





Db 156 ProSerSerProAlaProLeuProAlaProAlaSerProProSerProAlaProPro 175  
QY 54 CCACGAGCTCTCAAACTCGGCTCTTGAGACTCCGCATCCCCCGGCGCG 1  
Db 176 ProProAlaProProAlaPro-----HisProProSerProPro 189  
RESULT 3  
ADU02754  
ID ADU02754 standard; protein; 483 AA.  
XX ADU02754;  
AC ADU02754;  
XX 27-JAN-2005 (first entry)  
DT Novel human polypeptide seqid 1221.  
XX  
DE  
XX  
XX cytosolic; antiapoptotic; antiinflammatory; gene therapy; Nanodisc;  
KM proliferative disorder; inflammatory disorder; immune disorder;  
KM metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;  
KM ulcerative colitis; human.  
XX  
XX Homo sapiens.  
PN WO2004093804-A2.  
XX  
XX 04-NOV-2004.  
PD  
XX  
PF 19-APR-2004; 2004MO-US012047.  
XX  
XX 18-APR-2003; 2003US-0463708P.  
XX 18-APR-2003; 2003US-0463732P.  
PR 02-MAY-2003; 2003US-0467199P.  
PR 02-MAY-2003; 2003US-0467230P.  
PR 19-MAY-2003; 2003US-0471306P.  
PR 19-MAY-2003; 2003US-0471336P.  
PR 08-JUL-2003; 2003US-0485232P.  
PR 08-JUL-2003; 2003US-0485242P.  
PR 14-JUL-2003; 2003US-0486446P.  
PR 14-JUL-2003; 2003US-0486480P.  
PR 08-AUG-2003; 2003US-0493577P.  
PR 08-AUG-2003; 2003US-0493577P.  
PR 08-SEP-2003; 2003US-0505059P.  
XX  
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.  
PA  
XX Lee E, Hestir K, Chu K, Masuoka L, Williams LT;  
PI N-PSDB; ADU02022.  
DR WPI; 2004-775861/76.  
XX  
XX New first nucleic acid molecule comprising a polynucleotide sequence  
PT given in the specification, useful in preparing a composition for  
PT diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.  
XX  
XX Claim 14; SEQ ID NO 1221; 291pp; English.  
PS  
XX The invention describes a new first nucleic acid molecule comprising a  
CC polynucleotide sequence given in the specification. Also described are:  
CC an animal injected with the nucleic acid molecule; a second nucleic acid  
CC molecule comprising a second polynucleotide sequence that is at least  
CC about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or  
CC that hybridises to the first polynucleotide sequence under high  
CC stringency conditions; a vector comprising the nucleic acid molecule and  
CC a promoter that drives the expression of the nucleic acid molecule; a  
CC host cell transformed, transfected, transduced or infected with the  
CC nucleic acid molecule; a nucleic acid composition comprising a carrier or  
CC a buffer and one or more compositions comprising the nucleic acid  
CC molecule, vector or host cell; a substantially purified polypeptide; an  
CC animal injected with the polypeptide; a polypeptide composition  
CC comprising the polypeptide molecule and a carrier or buffer; a cell  
CC culture medium comprising the polypeptide or transfected cells  
CC transfected with the polynucleotide; making a transformed, transfected,  
CC transduced, or infected host cell; synthesising Nanodiscs simultaneously

CC and for synthesising a series of simultaneously-synthesised Nanodiscs  
CC sequentially utilising a dynamic system; preparing a hydrophobic protein  
CC for determination of crystal structure; immunising a non-human animal;  
CC screening for modulators of hydrophobic protein activity; a diagnostic  
CC kit; determining the presence of the nucleic acid molecule or its  
CC complement; determining the presence of an antibody to the polypeptide in  
CC a sample; an antibody specifically recognising, binding to or modulating  
CC the biological activity of at least one polypeptide encoded by a nucleic  
CC acid molecule or its biologically active fragment; an antibody  
CC composition comprising the antibody and a carrier; a bacteriophage, where  
CC the antibody is displayed on the bacteriophage; a bacterial cell  
CC comprising the bacteriophage; a non-human animal injected with the  
CC antibody composition; a host cell that secretes the antibody; making an  
CC antibody; diagnosing a disease, disorder, syndrome, or condition  
CC comprising cancer, or proliferative, inflammatory, immune, metabolic,  
CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or  
CC conditions in a patient; a modulator composition comprising a modulator  
CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a  
CC subject; an isolated modified cell comprising at least one first  
CC heterogeneous nucleic acid molecule, where the first heterogeneous nucleic  
CC acid molecule comprises a first polynucleotide sequence that encodes a  
CC first polypeptide; a non-human animal deficient in the polypeptide or  
CC that over-expresses the polypeptide; isolated tissues derived from the  
CC non-human animal; and one or more cells derived from the non-human  
CC animal. The nucleic acid is useful in preparing a composition for  
CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.  
CC This is the amino acid sequence of a novel human polypeptide of the  
CC invention.  
XX  
XX SQ Sequence 483 AA;  
XX  
XX Alignment Scores:  
XX Pred. No.: 7,49e-05 Length: 483  
XX Score: 142.50 Matches: 39  
XX Percent Similarity: 46.1% Conservative: 8  
XX Best Local Similarity: 38.2% Mismatches: 32  
XX Query Match: 15.4% Indels: 23  
XX DB: 8 Gaps: 4  
XX  
XX US-10-071-510A-16 (1-493) x ADU02754 (1-483)  
QY 273 CCATTTCTCCCGAATCCACTTCCTGCGATGCGAGTTTGTGATGAGCTCA 214  
Db 240 ProSerProSerProGlnProLeuProSerSerLeuProPro-----Ser 254  
QY 213 GCAGACTTTCTCCAGACCTCTCAAGCCCGCACGACTCTGCGAGCTCTGAGCTCT 154  
Db 255 SerProSerProProProProSerProSerSerSerLeuPro-----Pro 270  
QY 153 CCGTGCACACGCGACACCTCTCGGAGAAAGACTTTCATCAGACGACCGCTGCGCTT 94  
Db 271 ProAlaProProAlaProProProSerLeuProSerProProSerSerProLeu 290  
QY 93 CCACCGGAGACACT-----GGGCTCTCTTCT 67  
Db 291 ProProSerSerSerProSerProProProLeuSerAlaProProProSerProProSer 310  
QY 66 CCGGGAATTCGCGACACGACCTCTCAAACTCGGCTCTTGAGACTCGGATCCCGGCGC 7  
Db 311 ProSerProSerProProAlaPro-----ProLeuLeuSerProSerProProPro 327  
QY 6 CCGCGCG 1  
Db 328 ProPro 329  
RESULT 4  
ADT60739  
ID ADT60739 standard; protein; 238 AA.  
XX  
XX AC ADT60739;  
XX  
XX 13-JAN-2005 (first entry)  
DT  
XX

DE Plant polypeptide, SEQ ID 10816.  
XX Plant; transgenic; cold tolerance; growth rate; drought tolerance;  
KM disease resistance; galactomannan production; plant growth regulator;  
KM heat tolerance; herbicide tolerance; lignin production;  
KM extreme osmotic condition tolerance; pathogens resistance;  
KM pest resistance; yield improvement; seed oil yield; seed protein yield.  
XX  
XX Viridiplantae.  
XX  
XX US2004216190-A1.  
XX  
XX 28-OCT-2004.  
XX  
XX 18-DEC-2003; 2003US-00739930.  
XX  
XX 28-APR-2003; 2003US-00424599.  
XX 28-APR-2003; 2003US-00425115.  
XX  
XX (KOVA/) KOVALIC D K.  
XX  
XX KOVALIC DK;  
XX  
XX WPI; 2004-757369/74.  
XX  
XX New recombinant DNA constructs useful in the field of biochemistry and  
PT genetics, and in particular for producing transgenic plants with improved  
PT biological characteristics.  
XX  
XX Claim 2; SEQ ID NO 10816; 14pp; English.  
XX  
XX The invention relates a recombinant DNA construct comprising a  
XX polynucleotide having any of 5544 nucleotide sequences (CDNAs SEQ ID NO:  
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences  
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,  
CC Arabidopsis, wheat and rape but the specification does not indicate which  
CC sequences is derived from which organism. Also included is a method of  
CC producing a plant having an improved property, comprising transforming a  
CC plant with a recombinant DNA construct comprising a promoter region  
CC functional in a plant cell operably joined to a polynucleotide encoding a  
CC polypeptide associated with the property, and growing the transformed  
CC plant. The property is selected from improving plant cold tolerance, for  
CC manipulating growth rate in plant cells by modification of the cell cycle  
CC pathway, for improving plant drought tolerance, for providing increased  
CC resistance to plant disease, for galactomannan production, for production  
CC of plant growth regulators, for improving plant heat tolerance, for  
CC improving plant tolerance to herbicides, for increasing the rate of  
CC homologous recombination in plants, for lignin production, for improving  
CC plant tolerance to extreme osmotic conditions, for improving plant  
CC tolerance to pathogens or pests, for yield improvement by modification of  
CC photosynthesis, for modifying seed oil yield and/or content, for  
CC modifying seed protein yield and/or content, for yield improvement by  
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake  
CC and for yield improvement by providing improved plant growth and  
CC development under at least one stress condition. The polynucleotide may  
CC also encode a plant transcription factor. The methods and compositions of  
CC the present invention are useful in the field of biochemistry and  
CC genetics, in particular for producing transgenic plants with improved  
CC biological characteristics such as increased yield, improved nitrogen  
CC flow, increasing plant tolerance to cold or heat, improving plant  
CC tolerance to extreme osmotic and drought conditions, and improving plant  
CC tolerance to plant pests or pathogens. They can also be used in physical  
CC arrays of molecules, plant breeding markers, computer-based storage and  
CC analysis systems. The present sequence is one of the 5544 plant protein  
CC sequences of the invention. Note: The sequence data for this patent did  
CC not form part of the printed specification, but was obtained in  
CC electronic format directly from USPTO at  
CC [seqdata.uspto.gov/sequence.html?docId=20040216190](http://seqdata.uspto.gov/sequence.html?docId=20040216190).  
XX  
XX Sequence 238 AA;  
XX  
XX Alignment Scores: 8.3e-05 Length: 238  
XX  
XX Pred. No.:

Score: 141.00 Matches: 39  
Percent Similarity: 42.6% Conservative: 7  
Best Local Similarity: 36.1% Mismatches: 40  
Query Match: 15.3% Indels: 22  
DB: 8 Gaps: 4  
US-10-071-510A-16 (1-493) x ADP60739 (1-238)  
QY 273 CCATTTCTCCCGAATCCACTTCTCGAGATGCCAGTTTGTGAGGCTCA 214  
||| :||| ||| ||||| :|||  
DB 56 ProAlaGlyanProProProAlaProSerAlaProProAlaProThrProProGln 75  
QY 213 GCAGACTTCTTCCAGAGCC-----TCAGGCCCGCCAGCACTCTGCCA 169  
||| |||||  
DB 76 AlaProGlnAlaProAlaThrProProProAlaProGlnAlaProAlaThrProProPro 95  
QY 168 GCTCCCTAGCTCTCTCTGCA----- 148  
||| |||||  
DB 96 AlaPro\*\*ThrProProAlaThrProProProAlaProThrProProProAlaPro 115  
||| |||||  
QY 147 CCACGCGACACCCCTCCGAGAGACTTCTCCATCACCAGCCCTGCGCTTCCACCA 88  
||| ||| ||||| ||| :||| |||||  
DB 116 ProThrGlnProProProAlaProThrThrProProProSerProProAlaThrProPro 135  
QY 87 GGGACAGCTGG-----CTTCTTCTCGGGAATTGCAACGCTTCAAACTGG 34  
||| |||||  
DB 136 ProAlaProAlaThrProProProSerProProMetSerProProAlaThrProPro 155  
QY 33 CCTCTGGAGCTGGGCAATCCCG 10  
||| :||| |||||  
DB 156 PrometAlaThr-----ProPro 161  
RESULT 5  
ADN21309  
ID ADN21309 standard; protein; 639 AA.  
XX  
XX ADN21309;  
XX  
XX 02-DEC-2004 (first entry)  
XX  
XX Bacterial polypeptide #3962.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmotic;  
XX pathogens tolerance; pest tolerance; plant disease resistance;  
XX cell cycle pathway modification; plant growth regulator;  
XX homologous recombination; seed oil yield; protein yield; carbohydrate;  
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
XX bacterial polypeptide.  
XX  
XX Bacteria.  
XX  
XX US2003233675-A1.  
XX  
XX 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
XX (SLAT/) SLATER S C.  
XX (CHEN/) CHEN X.  
XX (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
PT for expression of a polynucleotide encoding a polypeptide from a  
PT microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 3962; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a  
 CC promoter functional in a plant cell, where the promoter is positioned to  
 CC provide for expression of a polynucleotide encoding a polypeptide from a  
 CC microbial source. The invention also relates to a transformed plant  
 CC comprising the recombinant DNA construct and a method of producing a  
 CC transformed plant having an improved property. The plant is a crop plant  
 CC such as maize or soybean. The method of producing a transformed plant  
 CC having an improved property comprises transforming a plant with the  
 CC recombinant DNA construct and growing the transformed plant, where the  
 CC polynucleotide or polypeptide is useful for improving plant properties.  
 CC The recombinant DNA construct is useful for producing plants with  
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,  
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
 CC increased resistance to plant disease, better growth rate by modification  
 CC of the cell cycle pathway with plant growth regulators, increased rate of  
 CC homologous recombination, modified seed oil or protein yield and/or  
 CC content, improved yield by modification of carbohydrate, nitrogen or  
 CC phosphorus use and/or uptake, by modification of photosynthesis or by  
 CC providing improved plant growth and development under at least one stress  
 CC condition, improved lignin production or improved galactomannan  
 CC production. This sequence represents a bacterial polypeptide used in the  
 CC scope of the invention. Note: The sequence data for this patent did not  
 CC form part of the printed specification but was obtained in electronic  
 CC format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 639 AA:

Alignment Scores:  
 Pred. No.: 0.000229 Length: 639  
 Score: 138.00 Matches: 39  
 Percent Similarity: 43.8% Conservative: 3  
 Best Local Similarity: 40.6% Mismatches: 36  
 Query Match: 15.0% Indels: 18  
 DB: 8 Gaps: 4

US-10-071-510A-16 (1-493) x ADN21309 (1-639)

QY 255 CCACCTTCACCTCTCTGACATGCCAGTTTCTGATGAGCTCAGACACTTCTT---CCA 199  
 Db 423 ProLeuProProLeuProAlaPro-----GlyProAlaProProLeuProPro 437  
 QY 198 GCAGCTTCAGAGCGCGCAGCACTCTGCGAGCTCCTGAGCTCTCTGACCAACGCGCAG 139  
 Db 438 AlaSerSerArgProProProMetLeuProThr-----ArgSerProAlaProProGln 455  
 QY 138 CACCTTCGCGAGAAAGACTTCTCCATCAGCAGCGAGCCCTGCGCTTCCA-----91  
 Db 456 AlaProProLeuProThrSerArgAlaProProProProProLeuProAlaThrGlnAla 475  
 QY 90 -----CCAGGACAGCTGGGACCTCTCTTCTCCGGAAATTCGCCACCA 49  
 Db 476 ProProProProProLeuProAlaThrSerAlaProProProProProProAlaProPro 495  
 QY 48 GCCTTCACAACTGGCCTTGGGACTCGGATCCCGCGGCGCGCG 1  
 Db 496 AlaProProAlaProProLeuProAlaAlaAlaAlaProProProProPro 511

RESULT 6  
 AAB74209  
 AAB74209 standard; protein; 731 AA.  
 AC AAB74209;  
 DT 17-MAY-2001 (first entry)  
 DE Protein encoded by Arabidopsis gene #4.  
 XX Proline-rich extensin-like receptor kinase; PERK; resistance; plant.  
 KW Arabidopsis thaliana.  
 OS

XX MO200114563-A1.

XX 01-MAR-2001.

XX 18-AUG-2000; 2000WO-CA000966.

XX 19-AUG-1999; 99US-0149466P.

XX 13-OCT-1999; 99US-0159122P.

XX (GORI/) GORING D.

XX (SILV/) SILVA N.

XX Goring D, Silva N;

XX WPI; 2001-244305/25.

XX New proline-rich, extensin-like receptor kinase nucleic acids and  
 PT polypeptides useful for increasing plant wounding or pathogen resistance,  
 PT or for producing transgenic plants with increased wounding or pathogen  
 PT resistance.

XX Example; Fig 14; 91pp; English.

XX The present invention relates to proline-rich extensin-like receptor  
 CC kinase (PERK). The PERK nucleic acids and polypeptides are useful for  
 CC increasing the resistance of plants to wounding and pathogens. These are  
 CC also useful for producing transgenic plants with increased wounding and  
 CC pathogen resistance compared with a wild type plant, as well as in assays  
 CC for identifying and developing compounds to inhibit and/or enhance  
 CC polypeptide function directly

XX Sequence 731 AA:

Alignment Scores:  
 Pred. No.: 0.000378 Length: 731  
 Score: 136.00 Matches: 52  
 Percent Similarity: 33.5% Conservative: 8  
 Best Local Similarity: 29.1% Mismatches: 54  
 Query Match: 14.7% Indels: 65  
 DB: 4 Gaps: 6

US-10-071-510A-16 (1-493) x AAB74209 (1-731)

QY 442 GCCCCAGACAGGAGGGGCTGTGCTGAGACCCCTACCC-----AC 401  
 Db 10 AlaProAlaThrSerProProAlaMetSerLeuPro-ProAlaSerValProAspThr 29  
 QY 400 CTCTTGACCTGGGGGCTGGCCCTTCTCCACAGCCCTGCTAGACAGACTCAGCGCTCGA 341  
 Db 29 rSerSerProProAlaProProLeuSerProLeuPro-----41  
 QY 340 CGATGCTGGGAATAGATCATGAGATTGATGAGAAATCTGACTTTGGATGTTGTG 281  
 Db 41 -----41  
 QY 280 GTGMAAACATTTCTTCCCGCAATCCACTTCATCTCTGCAGATGCCAGTTTGTGATG 221  
 Db 42 ----ProProLeuSerSerProProProLeuProSerProProProLeuSerAlaProTh 60  
 QY 220 AGGCTCAGCAGACTTCTTCCAGACGCTCAAGCGCGCAGAGACTGTGCGAGCTCC-- 163  
 Db 60 rAlaSerProProProLeuProValGluSerProProSerProProIleGluSerProPr 80  
 QY 162 -----TGAGCTCTCTGACCAACGCGACCACTT-----133  
 Db 80 oProProLeuLeuGluSerProProProProProLeuGluSerProSerProSerPr 100  
 QY 132 -----CCGGAGAAAGCTTCTCC 116  
 Db 100 oHisValSerAlaProSerGlySerProProLeuProPhleuProAlaLysProSerPr 120  
 QY 115 ATCACCAAGCAGGCCCT-----GCGCTTCCACAGGAGGACGCTGGGCTCTCTCC 65

Db	120	oProProSeSerProProSeSerCluThrvAlProFogIyAsnThrI]SeSerProPro	140
Qy	64	GGGAATTCTCCACACAGCCTCTCAACTCGGCTTGGAGCTGGCATCCCCG	10
Db	140	oArgSerLeuProSerCluSer---ThrProProValAsnThrAlaSerProPro	157
RESULT 7			
ABB93202	ID	ABB93202 standard; protein; 731 AA.	
AC	ABB93202;		
XX			
XX	31-MAY-2002	(first entry)	
XX			
XX	Herbicide; active polypeptide SEQ ID NO 2413.		
XX			
XX	Herbicide; plant; agriculture; herbicide.		
XX			
XX	Arabidopsis thaliana.		
XX			
XX	WO200210210-A2.		
XX			
XX	07-FEB-2002.		
XX			
XX	28-AUG-2001; 2001WO-EP009892.		
XX			
XX	28-AUG-2001; 2001WO-EP009892.		
XX			
XX	(FARB ) BAYER AG.		
XX			
XX	Tietjen K, Weidler M;		
XX			
XX	WPI; 2002-269010/31.		
XX			
XX	Identifying plant target proteins for herbicidally active compounds,		
XX	comprising aligning and comparing nucleic acid or amino acid sequences		
XX	from plant with nucleic acid or amino acid sequences from non-plant		
XX	organisms.		
XX			
XX	Claim 5; SEQ ID NO 2413; 261bp + Sequence Listing; English.		
XX			
XX	The invention relates to identifying target proteins (ABB90790-ABB94016)		
XX	for herbicidally active compounds, comprising aligning and comparing		
XX	nucleic acid or amino acid sequences from plant with nucleic acid or		
XX	amino acid sequences from non-plant organisms using suitable search		
XX	parameters, where plant sequences having an E-value greater by a factor		
XX	of 3 than the E-value of most similar non-plant sequences are selected.		
XX	The polypeptides or nucleic acids encoding them are useful for		
XX	identifying modulators. The identified modulators are useful as		
XX	herbicides		
XX			
XX	Sequence 731 AA;		
XX			
XX	Alignment Scores:		
XX	Pred. No.:	0.000378	Length: 731
XX	Score:	136.00	Matches: 52
XX	Percent Similarity:	33.5%	Conservative: 8
XX	Best Local Similarity:	29.1%	Mismatches: 54
XX	Query Match:	14.7%	Indels: 65
XX	DB:	5	Gaps: 6
XX			
XX	US-10-071-510A-16 (1-493) x ABB93202 (1-731)		
Qy	442	GGCCGAGACAGGAGGGGCTGTGCTGGAGAGACCCCTTACCC-----AC	401
Db	10	AlaProAlaThrSerProProAlaMetSerLeuPro-ProAlaAspSerValProAsp	29
Qy	400	CTTTGACCTGGGGGCTGGGCTTCTCCACACAGCCTGTAGACAGACTGCGTGA	341
Db	29	rSerSerProProAlaProProLeuSerProLeuPro-----	41
Qy	340	CGATGCTGGGAATAGCATTCATGGATTGATGAGAAATCTGACTTTGGGATGTTG	261

```

Db 41 ----- 41
Oy 280 GTGAAACCATTTTCTTCCCGAATTCACCTTCATCTCTGCAGATGCCAGTTTGTATG 22
    ||| ||||| ||||| ||||| |||||
Db 42 ----ProProluSerSerProProProluProSerProProProluSerAlaProth 60
Oy 220 AGGCTACGACGACTTCTTCCAGACGCCCTCAAGGCCCGCCACAGATCTGCCAGGTC-- 16
    ||| ||||| ||||| ||||| |||||
Db 60 rAlaSerProProProluProValGluSerProProSerProProlGluSerProPr 80
Oy 162 -----TGAGTCTCTCTGTCACACGGACGACCTT----- 133
    ||||| ||||| |||||
Db 80 oProProluLeuGluSerProProProProProluGluSerProSerProProSerPr 100
Oy 132 -----CCGGAAGAACTTCTCC 111
    ||||| ||||| |||||
Db 100 oHlSvAlSerAlaProSerGlySerProProProluProPheLeuProAlaLysProSerPr 120
Oy 115 ATGACCGACGACCGCCT-----GCGCTTCACACGAGGACAGCTGGGCTCTCTCC 65
    ||||| ||||| ||||| ||||| |||||
Db 120 oProProSerSerProProSerGluThrValProProGlyLanthrLieserProProPr 140
Oy 64 GGGAAATTCTGCGACGACCTCTCAAACTCGGCTCTTGAGGACTGGCATCCCCG 10
    ||||| ||||| ||||| ||||| |||||
Db 140 oArgSerLeuProSerGluSer---ThrProProValanThrAlaSerProPro 157

RESULT 8
ABR82942
ID ABR82942 standard; protein; 731 AA.
XX ABR82942;
XX AC
XX DT 18-DEC-2003 (first entry)
XX DB Arabidopsis PERK1 receptor related protein.
XX DE
XX KW PERK1; transgenic; plant; proline-rich extensin-like receptor kinase;
XX KM wound; pathogen resistance; plant growth; seed production.
XX OS Arabidopsis thaliana.
XX PN MO2003072763-A1.
XX PD 04-SEP-2003.
XX PE 28-FEB-2003; 2003MO-CA000274.
XX PR 28-FEB-2002; 2002CA-02373903.
XX PR 28-FEB-2002; 2002US-00086464.
XX PA (GORI/) GORING D.
XX PA (STLV/) SILVA N.
XX PA (HAF/) HAFRANI Y Z.
XX PI Goring D, Silva N, Haffani YZ;
XX DR WPI; 2003-712727/67.
XX DR N-PSDB; ACP36553.
XX PT Producing a transgenic plant having an increased plant resistance, plant
XX PT growth or seed production comprises transforming a plant with a nucleic
XX PT acid molecule having a Proline-rich Extensin-like Receptor Kinase
XX PT activity.
XX PS Disclosure; Fig 14; 123pp; English.
XX CC The invention relates to producing a transgenic plant having increased
XX CC plant height, number of branches, number of seed pods and/or seed
XX CC production compared to a non-transgenic plant, and/or quicker flowering
XX CC or later senescence compared to a non-transgenic plant. The method
XX CC involves transforming a plant with a vector including a Proline-rich
XX CC Extensin-like Receptor Kinase (PERK) nucleic acid molecule or a nucleic

```

CC acid molecule having PERK activity. The method, as well as the PERK  
CC nucleic acid molecule and polypeptide, are useful in increasing plant  
CC resistance to wounding and pathogens and in increasing plant growth and  
CC seed production. The nucleic acid molecule and polypeptide may also be  
CC used in producing transgenic plants or transgenic host cells. The present  
CC sequence represents a PERK1 polypeptide related protein from A. thaliana  
CC (Accession NO. CAA18590)

XX  
XX  
SQ Sequence 731 AA;

Alignment Scores:  
Pred. No.: 0.000378 Length: 731  
Score: 136.00 Matches: 52  
Percent Similarity: 33.5% Conservative: 8  
Best Local Similarity: 29.1% Mismatches: 54  
Query Match: 14.7% Indels: 65  
DB: 7 Gaps: 6

US-10-071-510A-16 (1-493) x ABR82942 (1-731)

QY 442 GCCCCAGAGGAGGCGCTGTGCTGAGACCCCTACCC-----AC 401  
DB 10 AAlProAlaThSerProProAlaMetSerLeuPro-ProAlaAspSerValProAspH 29  
QY 400 CTCTTGACCTGGGGGCGCTGCGCCCTTCTCCACAGCCCTGTGAGACAGACTCAGCGCTCA 341  
DB 29 fSerSerProProAlaProProLeuSerProLeuPro----- 41  
QY 340 CGATGCTTGGAATAGAGATCATGGATGATGAGAAATCTGACTTTGGAGTGTGTT 281  
DB 41 ----- 41  
QY 280 GTGAAACCATTTTCTTCCCGAATCCACTTCATCTCTGACAGATCCAGTTTCTGANG 221  
DB 42 ----ProProLeuSerSerProProLeuProSerProProProLeuSerAlaProH 60  
QY 220 AGGCTAGCAGACATTTTCTGACAGACCTTCAGAGCCCGCCAGCACTCTGCCAGCTCC-- 163  
DB 60 rAlaSerProProLeuProValAlaSerProProSerProProIleGluSerProPr 80  
QY 162 -----TGAGCTCTCTCTGACCAACGAGGAGACACCT----- 133  
DB 80 oProProLeuLeuGluSerProProProProProLeuGluSerProSerProSerPr 100  
QY 132 -----CCGAGAAAGACTTCTCC 116  
DB 100 oHisValSerAlaProSerGlySerProProLeuProHleuProAlaAlaSerProSerPr 120  
QY 115 ATCACCAGCAGGCGCT-----GCGCTTCCACAGGAGGAGCTGGGCTTCTCTCC 65  
DB 120 oProProSerSerProProSerGluThrValProProGlyAlaThrIleSerProProPr 140  
QY 64 GGAATTTCTGCCACAGCCTCTCAAACTCGGCTCTTGGGACTCGGATCCCCG 10  
DB 140 oArgSerLeuProSerGluSer---ThrProValAlaThrAlaSerProPro 157

RESULT 9

ABG27250  
ID ABG27250 standard; protein; 406 AA.

XX  
XX  
AC ABG27250;

XX  
XX  
DT 18-FEB-2002 (first entry)

XX  
XX  
DE Novel human diagnostic protein #27241.

XX  
XX  
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX  
XX  
PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX  
XX  
PF 30-MAR-2001; 2001WO-US008631.

XX  
XX  
PR 31-MAR-2000; 2000US-00540217.

XX  
XX  
PR 23-AUG-2000; 2000US-00649167.

XX  
XX  
PA (HISE-) HISEQ INC.

XX  
XX  
PI Drmanac RT, Liu C, Tang YT;

XX  
XX  
DR MPI: 2001-639362/73.

XX  
XX  
DR N-PSDB; AAS91437.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX  
XX  
PS Claim 20; SEQ ID NO 57609; 103pp; English.

XX  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostic as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful for treating disorders  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABE00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
XX  
SQ Sequence 406 AA;

Alignment Scores:  
Pred. No.: 0.000614 Length: 406  
Score: 133.00 Matches: 35  
Percent Similarity: 39.8% Conservative: 4  
Best Local Similarity: 35.7% Mismatches: 47  
Query Match: 14.4% Indels: 12  
DB: 4 Gaps: 3

US-10-071-510A-16 (1-493) x ABG27250 (1-406)

QY 300 CTGACTTTGGAGATGTTGTTGTAACATTTTCTTCCCGAATCCACTTCATCCTCT 241  
DB 13 LeuLeuLeuGlnGlySerLeuProProSerProLeuProSerProProSerProSerPro 32  
QY 240 GCAGATGCCAGCTTCTGATGAGGCTCAGCAGACTTCTTCCAGCAGACCTCAAGCCGCC 181  
DB 33 ProProProPro-----ProProProProSerProSer 43  
QY 180 ACAGACTCTGCCAGCTCTGAGCTCTCTGACCAACGAGAGACCTCCGAGAAAGACT 121  
DB 44 SerProSerProProProSerProProProProProProSerProSerProSerPro 63  
QY 120 TCTTCATCAGCAGCAGCTCTGCGCTTCCACAGGAGAGACTGGGCT---CCTTCGCG 64  
DB 64 SerProSerProProProProProThrSerProProSerProSerProSerProSerPro 83  
QY 63 GGAATTTCTGCCACAGCCTCTCAAACTCGGCTCTTGGGACTCGGATCCCCCG 10  
DB 63 GGAATTTCTGCCACAGCCTCTCAAACTCGGCTCTTGGGACTCGGATCCCCCG 10  
XX  
XX  
PN WO200175067-A2.

```

Db          84 ProProSerProProSerProSer-----ProLeuSerProThrProProPro 99
RESULT 10
ABG14000
ID   ABG14000 standard; protein, 598 AA.
XX
AC   ABG14000;
XX
DT   18-FEB-2002 (first entry)
XX
DE   Novel human diagnostic protein #13991.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX   Food supplement; medical imaging; diagnostic; genetic disorder.
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US008631.
XX
PR   31-MAR-2000; 2000US-00540217.
XX   23-AUG-2000; 2000US-00649167.
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI, 2001-639362/73.
XX
DR   N-PSDB; AAS78187.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity.
XX
PS   Claim 20; SEQ ID NO 44359; 103bp; English.
XX
XX
CC   The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC   sequences. (I) is useful as hybridisation probes, polymerase chain
CC   reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC   and in recombinant production of (II). The polynucleotides are also used
CC   in diagnostics as expressed sequence tags for identifying expressed
CC   genes. (II) is useful in gene therapy techniques to restore normal
CC   activity of (II) or to treat disease states involving (II). (II) is
CC   useful for generating antibodies against it, detecting or quantitating a
CC   polypeptide in tissue, as molecular weight markers and as a food
CC   supplement. (II) and its binding partners are useful in medical imaging
CC   of sites expressing (II). (I) and (II) are useful for treating disorders
CC   involving aberrant protein expression or biological activity. The
CC   polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependant on DNA and
CC   amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic
CC   amino acid sequences of the invention. Note: The sequence data for this
CC   patent did not appear in the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 598 AA;

```

```

Alignment Scores:
Pred. No.:      0.000699      Length:      598
Score:          133.00        Matches:      40
Percent Similarity: 34.0%      Conservative: 8
Best Local Similarity: 28.4%    Mismatches:  37
Query Match:    14.4%         Indels:       56
DB:             4            Gaps:          5

```

```

US-10-071-510A-16 (1-493) X ABG14000 (1-598)

```

```

OY          411 CCCCTACCCACCTTTGACCTGGGCGCTGGCCCTTCTCCGACAGCCCTGTAGACAGAC 352
Db          130 ProLeuProSerProSerProSerAlaLeProSerLeuProProPro----- 145
OY          351 TCACGCGTCAGCATGCTGGGAATGAGATCCATGGATTGATGAGAAATCCTGACTTTG 292
Db          145 ----- 145
OY          291 GGATGTTGTTGGTGAANAACATTTTCTTCCCGAATCCACTTCCTGTGAGATGCC 232
Db          146 -----SerProGlnProLeuProPro----- 152
OY          231 AGTTTGTGATGAGCTCAGACACTTTCTTCAGCAGCGCTCAAGGCCCGCAGACACTCTG 172
Db          153 -----ProProProSerSerProProProSerLeu 162
OY          171 CCAGCTCCCTGAGCTCTCTCTGCACCA--CGCAGCACCTCCGAGAGACTTTCTCA 115
Db          163 ProSerProLeuLeuProProProProLeuSerSerSerSerProLeuSerPro 182
OY          114 TCACGAGCGACCCCTGCGCCTTCCACCAAGG-----ACAGCTGGGCT 73
Db          183 SerProProProProSerProProProSerLeuProProSerProProProSerProPro 202
OY          72 CCTTCTCCGGAATTTCTGCCACACGACCTCTCAAACTTCGCGCTTTGGAGACTCGGCATCC 13
Db          203 ProProProProProGlnProProSerProProSerProProSerSerProLeuSerSer-----Pro 220
OY          12 CCG 10
Db          221 Pro 221

```

```

RESULT 11
ADP60423
ID   ADP60423 standard; protein, 598 AA.
XX
AC   ADP60423;
XX
DT   12-FEB-2004 (first entry)
XX
DE   Human contig polypeptide sequence SEQ ID NO:2790.
XX
KW   biological activity; genetic engineering; hybridisation probe; oligomer;
KW   primer; chromosome mapping; gene mapping; recombinant protein production;
KW   human.
XX
OS   Homo sapiens.
XX
PN   WO2003080795-A2.
XX
PD   02-OCT-2003.
XX
PF   09-AUG-2002; 2002WO-US025485.
XX
PR   09-AUG-2001; 2001US-0311261P.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Yang Y, Wang Z, Wang G, Ma Y;
XX
DR   WPI, 2003-876918/81.
XX
DR   N-PSDB; ADP59971.
XX
PT   New polynucleotides, useful as hybridization probes, oligomers or
PT   primers, for chromosome or gene mapping, for the recombinant production
PT   of proteins, and for generating antisense DNA or RNA.
XX
PS   Example 3; SEQ ID NO 2790; 571bp; English.
XX
CC   The present invention describes isolated polynucleotide sequences (I),
CC   which encode polypeptides (II) with biological activity. Also described:
CC   (1) a vector comprising (I); (2) an expression vector comprising (1); (3)

```





```

Db 96 -----ProPterSerSerThrProProProThnLeuProProProSerProProAsp 112
Qy 224 GCCAGTTTCTGATGAGCTCAGCAGACTTTCTTCAGACGCTCA-----AGCCCGCC 181
Db 113 -----SerProProLeuProProAlaSerProThrProProVal 126
Qy 180 AGCAGCTGCGCA-----GCTCCCTGAGCTCTCTCCACACCGG 142
Db 127 ThnThrSerProSerProProValThnThrSerProSerProSerProAlaThrThr 146
Qy 141 CAGCACCCTCCGAGAGACTTCTCCATCA-----CCAGCCAGCCT 100
Db 147 SerProProProAlaGlnThrProProSerLeuProAsnLeuSerProProSerProPro 166
Qy 99 GCG-----CTTCACAGGAGAGACGCTGGCCTCT-----70
Db 167 AlaGlySerProProProGlnSerProProThnThrLeuProProProSerArgProLe 186
Qy 69 TCTCCGGGAATTCGCCAGCAGCTCTCAACTCGGCGCTTGGAGCTCGGATCCCGG 10
Db 187 SerProSerProProProAlaAsnProAlaArgProProThrThrArgThrProPro 206

RESULT 13
AAW31852
ID AAW31852 standard; protein; 763 AA.
XX
AC AAW31852;
XX
DT 25-MAR-2003 (revised)
DT 27-APR-1998 (first entry)
XX
DE Mycobacterium tuberculosis 74 kDa protein.
XX
KM Tuberculosis; mycobacteria; infection; diagnosis; antimycobacterial;
XX antibiotic; vaccine.
XX
OS Mycobacterium tuberculosis.
XX
FN W09741252-A2.
XX
PD 06-NOV-1997.
XX
PF 18-APR-1997; 97MO-EP001973.
XX
PR 29-APR-1996; 96DE-01017184.
XX
PA (GBFB ) GBF GBS BIOTECH FORSCHUNG GMBH.
XX
PI Singh M, Honisch C, Esplicita C, Moreno C;
XX
DR WPI; 1997-549750/50.
XX
DR N-PSDB; AAT93610.
XX
PT New DNA and related proteins or RNA derived from M. tuberculosis - used
PT for diagnosis of mycobacterial infections, monitoring vaccination and
PT development of anti-mycobacterial agents.
XX
PS Claim 5; Fig 13; 55pp; English.
XX
CC This novel 74 kDa protein is encoded by an open reading frame of a
CC Mycobacterium tuberculosis DNA fragment (see AAT93610) containing
CC polymorphic GC-rich sequences. Its amino acid sequence shows a high
CC proline content, but there is no homology to any known proline-rich
CC antigens of mycobacteria. Novel M. tuberculosis proteins (see AAW31851-
CC 57) are claimed. These can be produced as recombinant proteins,
CC especially in bacterial, yeast, fungal or higher eukaryote host cells,
CC and used for diagnosing tuberculosis and other mycobacterial infections
CC in humans or animals. The claimed proteins can also be used for
CC epidemiological studies, for monitoring vaccination, and for the
CC development of vaccines and anti-mycobacterial drugs. (Updated on 25-MAR-
CC 2003 to correct PR field.)
XX

```

```

SQ Sequence 763 AA;
Alignment Scores:
Pred. No.: 0.00107 Length: 763
Score: 131.50 Matches: 47
Percent Similarity: 32.9% Conservative: 10
Best Local Similarity: 27.9% Mismatches: 53
Query Match: 14.2% Gaps: 64
DB: 2
US-10-071-510A-16 (1-493) x AAW31852 (1-763)
Qy 410 CCTACCCAGCTCTTGAGCTGGGCT-----GGCCCTCTCCACAGC 366
Db 151 ProPheProProAlaAlaLeuAsnProProAlaProProAlaProProLeuAlaAsnPro 170
Qy 365 -----CCTGTAAGACAGACTGACGCGTGCAGATGCTGGGAATAGATCCAT 318
Db 171 ProProLeuProProAlaProProThrProAlaGlyThrProProAlaAlaPro----- 188
Qy 317 GGGATTGATGAGAAATCCTGACTTTGGGATGTTGTGTGA-----AAAGCA 271
Db 189 -----TTPProProValProAlaAlaProLysSerLysPro 200
Qy 270 TTTTCTCCCGAATCCACTTCATCCTGTCAGATGCCAGTTTCTGATGAGCTCAGCA 211
Db 201 AlaSerProProAlaArgPro-----206
Qy 210 GACTTTCTTCAGACAGCTCAAGCGCCGACAGACTGCGCACTCCCTGAGCTCTCT 151
Db 207 -----ProAlaProProMetProAlaThrProMetGluPheProProLeuProPro 223
Qy 150 GCACCAAGGAGACACCTCCGAGAGACTTCCATGACAGCAGCCGCTGGCTTCA 91
Db 224 ValProProAspProLysSerLysGlnThrProProAlaProProAlaProProLysPro 243
Qy 90 CCAGGAGACAGCTGGG-----CCTCTCTCCGGAATTCGCA 52
Db 244 ProAlaProProAlaProProValAlaValAlaValAlaProCysArgProPro 283
Qy 51 CCAGCTCTCAAACT-----37
Db 264 ProAlaProProAlaProProValAlaValAlaValAlaProCysArgProPro 283
Qy 36 ---CGGCTCTTGAGACTCGGATCCCGGCGCGCGG 1
Db 284 LeuProProLeuProProAsnAsnHisProProAlaProPro 296

RESULT 14
ABB91504
ID ABB91504 standard; protein; 708 AA.
XX
AC ABB91504;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 715.
XX
KM Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
FN W0200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX

```



PR	06-JUL-1999	99US-01420555
PR	06-JUL-1999	99US-0142380P
PR	08-JUL-1999	99US-0142803P
PR	09-JUL-1999	99US-0142920P
PR	12-JUL-1999	99US-0142977P
PR	13-JUL-1999	99US-0143542P
PR	14-JUL-1999	99US-0143654P
PR	15-JUL-1999	99US-0144005P
PR	16-JUL-1999	99US-0144085P
PR	16-JUL-1999	99US-0144086P
PR	19-JUL-1999	99US-0144335P
PR	19-JUL-1999	99US-0144331P
PR	19-JUL-1999	99US-0144332P
PR	19-JUL-1999	99US-0144333P
PR	19-JUL-1999	99US-0144334P
PR	19-JUL-1999	99US-0144335P
PR	19-JUL-1999	99US-0144336P
PR	20-JUL-1999	99US-0144352P
PR	20-JUL-1999	99US-0144652P
PR	20-JUL-1999	99US-0144684P
PR	21-JUL-1999	99US-0144814P
PR	21-JUL-1999	99US-0145086P
PR	22-JUL-1999	99US-0145088P
PR	22-JUL-1999	99US-0145089P
PR	22-JUL-1999	99US-0145097P
PR	22-JUL-1999	99US-0145192P
PR	23-JUL-1999	99US-0145145P
PR	23-JUL-1999	99US-0145218P
PR	23-JUL-1999	99US-0145224P
PR	26-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145913P
PR	27-JUL-1999	99US-0145918P
PR	27-JUL-1999	99US-0145919P
PR	28-JUL-1999	99US-0145951P
PR	02-AUG-1999	99US-0146386P
PR	02-AUG-1999	99US-0146388P
PR	02-AUG-1999	99US-0146389P
PR	03-AUG-1999	99US-0147038P
PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147302P
PR	05-AUG-1999	99US-0147192P
PR	05-AUG-1999	99US-0147260P
PR	06-AUG-1999	99US-0147303P
PR	06-AUG-1999	99US-0147416P
PR	09-AUG-1999	99US-0147493P
PR	09-AUG-1999	99US-0147935P
PR	10-AUG-1999	99US-0148171P
PR	11-AUG-1999	99US-0148319P
PR	12-AUG-1999	99US-0148341P
PR	13-AUG-1999	99US-0148565P
PR	13-AUG-1999	99US-0148684P
PR	16-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-0149175P
PR	18-AUG-1999	99US-0149426P
PR	20-AUG-1999	99US-0149722P
PR	20-AUG-1999	99US-0149723P
PR	20-AUG-1999	99US-0149929P
PR	22-AUG-1999	99US-0149902P
PR	23-AUG-1999	99US-0149930P
PR	23-AUG-1999	99US-0150586P
PR	26-AUG-1999	99US-0150884P
PR	27-AUG-1999	99US-0151065P
PR	27-AUG-1999	99US-0151066P
PR	27-AUG-1999	99US-0151080P
PR	30-AUG-1999	99US-0151303P
PR	31-AUG-1999	99US-0151438P
PR	01-SEP-1999	99US-0151930P
PR	07-SEP-1999	99US-0152263P
PR	10-SEP-1999	99US-0153070P
PR	13-SEP-1999	99US-0153758P
PR	15-SEP-1999	99US-0154018P
PR	16-SEP-1999	99US-0154039P
PR	20-SEP-1999	99US-0154779P

PR	23-SEP-1999;	99US-01551399;
PR	23-SEP-1999;	99US-0155468P
PR	24-SEP-1999;	99US-0155659P
PR	28-SEP-1999;	99US-0156458P
PR	29-SEP-1999;	99US-0156569P
PR	04-OCT-1999;	99US-0157117P
PR	05-OCT-1999;	99US-0157753P
PR	06-OCT-1999;	99US-0157865P
PR	07-OCT-1999;	99US-0158029P
PR	08-OCT-1999;	99US-0158282P
PR	12-OCT-1999;	99US-0158369P
PR	13-OCT-1999;	99US-0159253P
PR	13-OCT-1999;	99US-0159284P
PR	13-OCT-1999;	99US-0159285P
PR	13-OCT-1999;	99US-0159293P
PR	14-OCT-1999;	99US-0159330P
PR	14-OCT-1999;	99US-0159331P
PR	14-OCT-1999;	99US-0159657P
PR	14-OCT-1999;	99US-0159658P
PR	18-OCT-1999;	99US-0159584P
PR	21-OCT-1999;	99US-0160741P
PR	21-OCT-1999;	99US-0160767P
PR	21-OCT-1999;	99US-0160768P
PR	21-OCT-1999;	99US-0160774P
PR	21-OCT-1999;	99US-0160814P
PR	21-OCT-1999;	99US-0160815P
PR	22-OCT-1999;	99US-0160980P
PR	22-OCT-1999;	99US-0160981P
PR	22-OCT-1999;	99US-0160989P
PR	25-OCT-1999;	99US-0161404P
PR	25-OCT-1999;	99US-0161405P
PR	25-OCT-1999;	99US-0161406P
PR	26-OCT-1999;	99US-0161359P
PR	26-OCT-1999;	99US-0161360P
PR	26-OCT-1999;	99US-0161361P
PR	28-OCT-1999;	99US-0161920P
PR	28-OCT-1999;	99US-0161922P
PR	28-OCT-1999;	99US-0161933P
PR	29-OCT-1999;	99US-0162142P

Alignment Scores:	
Pred. No.:	0 000942
Score:	130.00
Percent Similarity:	36.7%
Best Local Similarity:	31.7%
Query Match:	14.1%
DB:	3
US-10-071-510A-16 (1-493) x AAC29836 (1-191)	
	length: 191
	Matches: 44
	Conservative: 7
	Mismatches: 38
	Indels: 50
	Gaps: 7

Qy 57 CTGCCACGAGCTCTCAAACTCGGCTTTGGGACTCGGCATCCCCGGGCGCG 1  
|||  
::: |||  
Db 99 SerProProProAlaThrProProProValAlaThr-----ProProProAlaPro 115

Search completed: April 29, 2006, 03:41:52  
Job time : 142 secs

**THIS PAGE LEFT BLANK**



Db	41	-----	41
Qy	280	GTGAACCATTTTCTTCCCGAATCCACTTCACCTCTGACAGATGCAGTTCATG	221
Db	42	---ProProuserSerProProProuserSerProProProuserAlaProTh	60
Qy	220	AGGCTCACAGACTTCTTCAGCAGCCTCAAGGGCCGACAGACTGACAGCTCC	163
Db	60	rAlaserProProProuserProValGluSerProProProuserProProuserProPr	80
Qy	162	-----TGAGCTCCTCTGCACACGAGCAGCACTT	133
Db	80	oProProLeuLeuGluSerProProProProProuLeuGluSerProSerProProSerPr	100
Qy	132	-----	116
Db	100	oH1eValSerAlaProSerGlySerProProProuProPheLeuProAlaLysProSerPr	120
Qy	115	ATCACCGACGACCTT	65
Db	120	oProProSerSerProProSerGluThValProProGlyLysnThrIleSerProProPr	140
Qy	64	GGGAATTTGCAACAGCCTCTCAAACTGGGCTTTGAGCTGGCATCCCCG	10
Db	140	oArgSerLeuProSerGluSer---ThrProProValnThrAlaSerProPro	157

```

Oy      42 CAAACCTGCG-----CCTCTGGGACCTCCGCATCCCCGGGCGCG 1
Db      415 ArgSerArGAlaGlyGlyProPfoLeuGLYThrArGProPfoProPfoPro 432

RESULT 3
T29293
hypochemical protein C50F7.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29293
R:Johnson, D.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C50F7.
A:Reference number: Z20601
A:Accession: T29293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-309 <UOH>
A:Cross-references: UNIPROT:Q18751; UNIPARC:UPI0000060F74; EMBL:U41557; PIDN:AAA83301.1
C:Genetics:
A:Gene: CBSP:C50F7.5

Alignment Scores:
Pred. No.:      0.00408      Length:      309
Score:          133.50      Matches:      52
Percent Similarity: 36.8%      Conservative: 8

```

```

RESULT 2
S22697
extension - Volvox carteri (Fragment)
C:Species: Volvox carteri
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S22697, S21006
C:Ercl, H.; Hallmann, A.; Wenzl, S.; Sumper, M.
EMBO J. 11, 2055-2062, 1992
A:Title: A novel extension that may organise extracellular matrix biogenesis in Volvox carteri
A:Reference number: S22697, MUID:92289669, PMID:1600938
A:Accession: S22697
A:Molecule type: mRNA
A:Residues: 1-464 <HAL>
A:Cross-references: UNIPROT:Q41645; UNIPARC:UPI000009F7C8; EMBL:X65165; NID:G21991; PIDN:
C:Keywords: glycoprotein

Alignment Scores:
Pred. No.: 0 00348 Length: 464
Score: 134.50 Matches: 40
Percent Similarity: 40.7% Conservative: 8
Best Local Similarity: 33.9% Mismatches: 33
Query Match: 14.6% Indels: 37
DB: 2 Gaps: 5

US-10-071-510A-16 (1-493) x S22697 (1-464)

QY 273 CCAATTTCTCCCGGAATCCATCTTCATCTCGATGCCAGTTTCTGATGAGCTCA 214
||| ||||| ||||| ::||| ||| |||
Db 325 ProArgSerSerProSerProProProSerProProProPro----- 339
213 GCAGACTTCTTCCAGCAGCCTCAAGCCGCGCAGCATCTGCAGTCCCTGAGCT--- 157
||| ||| ||||| ::||| ::||| ||| ::|
Db 340 -----SerProProProProArgProSerProSerProProProProArgSerSer 356
156 -----CCTCTGCACACGCGCAGCACCCCTCG 130
||| ||||| ||||| ||||| |||||
Db 357 ProSerProProProProProValSerProProProProProProArgSerProPro 376
129 GAGAAAGATTCCTCCATCAACGACCGCCCTGGCTTCCACGAGGAGACGAGGGCTGCT 70
::: ||||| ||| ||| ||||| ::| |||||
Db 377 ProProProAlaSerSerProProProProProArgProProProProProProPro 396
69 TCTCCGGGAATTCGCGCACAGCC-----TCT 43
||||| ||||| ||||| |||||
Db 397 SerProPro-----ProProAlaThrAlaAlaAlaSerProProSerProAlaProSer 414

```

```

Query Match:      14.5%      Indels:      42
DB:               2         Gaps:          7

US-10-071-510A-16 (1-493) x T29293 (1-309)

OY      467 CCAAGTCTCTAGACGACGATGTTGCCCCAGACGAGGAGGCGCTTCTCGAGACCCC 408
      ||| |||
      ::| |||
Db      151 ProGlyProValAspProSerGluAspProGlnProSerValGluProSerGluAspPro 170
      ||| |||
OY      407 TACCACCTCTTGACCTGTGGGCGCTGGCCCTCTTCCACAGCCCTGCTAGACATCTAC 348
      ||| |||
      ||| |||
Db      171 GlnProSerGlyProProSerProGlyProValAspProSerGluAspProGlnProSer 190
      ||| |||
OY      347 GCGTCGACGATCCTGTGGAAATAG-GATTCATGGGATTGATGAGAAATCTGACTTTGGGA 289
      ||| |||
      ||| |||
Db      191 GlySerSerSerProGlyProValAspProSerAsp----- 202
      ||| |||
OY      288 TGTGTTGTGGAAACCATTTTCTCCCGCAATCTTCATCTCTGCGAGATCCAGT 229
      ||| |||
      ||| |||
Db      203 -----GluProSerProSerGlyProProSerProGlyProValAspProSer 218
      ||| |||
OY      228 TTCTGATGAGGCTTCAGACGACTTCTTCAGACGCTCAAGGCCGCCAGACTTCGCCA 169
      ||| |||
      ||| |||
Db      219 -----GluAspPro 221
      ||| |||
OY      168 GCTCCCTGAGCTCTCTCTGACCA CGGCGACGACCCCTCGGAGAACTTCTCATCAACA 109
      ||| |||
      ||| |||
Db      222 LysProSerGluProProSerProGlyProValAspPro--SerAspGluProSerPro 240
      ||| |||
OY      108 GCCAGCCCTCGCGCTTCACACGAGGACGACTGGGCTCTCTCT-----CCGGGAATT 58
      ||| |||
      ||| |||
Db      241 SerAspProProGlyProProGlyProProGlyProProProThrArgArgProProGly--- 259
      ||| |||
      ||| |||
OY      57 CTGGCACCAGCCTCTCAACTCG-----CCTCTTGGG-----ACTCGGCAT 16
      ||| |||
      ||| |||
Db      260 ---ProProGlyProProThrArgArgProProGlyProProGlyProProThrArgArg 278
      ||| |||
OY      15 CCCCCGGGC 7
      ||| |||
      ||| |||
Db      279 ProProGly 281

RESULT 4
D96711
hypochemical protein P24J5.8 [imported] - Arabidopsis thaliana
[CSpecies: Arabidopsis thaliana (mouse-ear cress)]
[Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text change 09-Jul-2004]

```



C:Accession: D96711  
 R:Theologian, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: D96711  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-708 <STO>  
 A:Cross-references: UNIPROT:Q95XJ1; UNIPARC:UP10000048426; GB:AE005173; NID:G5734709; PI  
 C:Genetics:  
 A:Map position: 1  
 A:Map position: 1

Alignment Scores:  
 Score: NO.: 0.00764 Length: 708  
 Percent Similarity: 130.50 Matches: 37  
 Best Local Similarity: 41.1% Conservative: 9  
 Query Match: 33.0% Mismatches: 39  
 DB: 14.1% Indels: 27  
 Gaps: 5

US-10-071-510A-16 (1-493) x D96711 (1-708)

Qy	273	CCATTTCCTTCCTCCG-----AATCCACTTTCCTTCGAGATGCCAGTTCTGA	223
Db	29	ProAlatnProProProProValThnSerProLeuProProSerAlaPro-----	44
Qy	222	TGAGGCTCAGACGACTTCTTCGACGACCTTCAGAGCCCGCCGACGACTCTCCAGCTCCC	163
Db	45	-----ProProAlaValAlaProProProProProProValThnThnSerProPro	62
Qy	162	-----TGAGTCTCTCTGCACACGCGACGACCCCTCCGAGAACTTCT-----	118
Db	63	ValAlaAlaenglyAlaProProProProProLeuProlyspProGluSerSerProPro	82
Qy	117	CCATCACCAGCAGCCCTGCGCTTCACACGAGGACAGCTGGGCTCTCTCT-----	67
Db	83	ProGlnProValIleProSerProProSerThnSerProProGlnProValIle	102
Qy	66	-----CCGGGAATTCGCGACGAGCCTCTCAACT	37
Db	103	ProSerProProProSerAlaSerProProProAlaValProProLeuProSerSer	122
Qy	36	CGGCTCTTGGGACTCGGCATCCCGCGGCGCGCCG	1
Db	123	ProProProProAlaSerValProProProAlaGPro	134

RESULT 5  
 F84522  
 Probable proline-rich protein [imported] - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 C:Accession: F84522  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Unayam, L.; Tallon,  
 usen, D.; Nierman, W.C.; White, O.; Eiseen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: F84522  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-191 <STO>  
 A:Cross-references: UNIPROT:082327; UNIPARC:UP1000009F0P1; GB:AE002093; NID:G3650031; P

A:Gene: At2g14890  
 A:Map position: 2  
 C:Superfamily: proline-rich protein 3

**Alignment Scores:**  
 Pred. NO.: 0.00763 Length: 191  
 Score: 130.00 Matches: 44  
 Percent Similarity: 36.7% Conservative: 7  
 Best Local Similarity: 31.7% Mismatches: 38  
 Query Match: 14.1% Indels: 50  
 DB: 2 Gaps: 7

US-10-071-510A-16 (1-493) x F84522 (1-191)

QY	405	CCCACTCTTGACCTGGGGCTGGCCCTCTCCACAGCCCTCTAGACAGACTCAGC	346
DB	23	ProThrSerProProThrAlaThrProAlaProProThrProThr-----	37
QY	345	GTCGACGATGCTGGGATAGATCCATGGATGTATGAGAAATCTGACTTTGGGATGT	286
DB	38	-----ThrPro-----	39
QY	285	TGTTGGTGAACCATTTTCTTCCCGAATCCACTTCATCTCTGACAGATGCAGTTTC	226
DB	40	-----ProProAlaAlaThrProProProVal-----	48
QY	225	TGATGAGCTCAGCAGACTTCTTCCAGCAGCCCGACAGCCCGCAGACTTGGCAGCT	166
DB	49	-----SerAlaProProProValThrThrSerProPro	59
QY	165	CCC-----TGAGCTCTCTCGCAGCAGCAGCAGCCTCCGAGAGACTTTCATCA	112
DB	60	ProValThrThrAlaProPro-----ProAlaAsnProProProProValSerSerProPro	78
QY	111	CCAGCCAGC-----CTTGGCTTCCACCGAGGACAGCTGGGCTCTTCTCCGGGAAT	58
DB	79	ProAlaSerProProProAlaThrProProProProValAlaSerProProProProValAla	98
QY	57	CTGCCACCGACCTCTCAAACTCGGCTTGGGACTGGAGTCCCGGGCCGGCCG	1
DB	99	SerProProProAlaThrProProProProValAlaThr-----ProProProAlaPro	115

**RESULT 6**  
 T04859  
 extensin homolog F28A21.80 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 C:Accession: T04859  
 R:Bevan, M.; Mueller, M.W.; Mundlein, A.; Felber, R.; Baneroff, I.; Mewes, H.W.; Meyer  
 submitted to the Protein Sequence Database, February 1999  
 A:Reference number: Z15387  
 A:Accession: T04859  
 A:Molecule type: DNA  
 A:Residues: 1-839 <BEV>  
 A:Cross-references: UNIPROT:O9SN46; UNIPARC:UPI00000A17EF; EMBL:AL035526  
 A:Experimental source: cultivar Columbia; BAC clone F28A21  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 623/3  
 A>Note: F28A21.80

**Alignment Scores:**  
 Pred. No.: 0.0103 Length: 839  
 Score: 129.00 Matches: 49  
 Percent Similarity: 40.8% Conservative: 20  
 Best Local Similarity: 29.0% Mismatches: 54  
 Query Match: 14.0% Indels: 46  
 DB: 2 Gaps: 7

US-10-071-510A-16 (1-493) x T04859 (1-839)

QY	404	CCACCTCTTGACCTGGG-----CCTGGC-----	381

Db 356 proProValcylLeuGlyLeuProGlyLeuPheAspAaArgArgAsnCylLeuProIlaArg 375  
 QY 380 CCCTTCTCCACAGAGCCCTGCTAGACAGACTCACGGCTGCAGATGCCCTGGAAAT- 326  
 Db 376 ProIlaGlnArgSerProGlyGlnCysAlaAlaPheSerSerLeuPro-ProValAspCyl 395  
 QY 325 -GGATTCATGGAGATTGATGAGAAATCTGCAGCTTTGGAGTGTGTGTGGAAACCATTTT 267  
 Db 395 sGlySerheGlyCysGlyArgSerThrArgProProValValPro-SerProProt 415  
 QY 266 CTTCGCCGAATCCCATTCATCCTCTGCAGATGCCAGTTTCTGATAGAGCTCAGACAGT 207  
 Db 415 hTrhProSerPro-----GlyGlySerPro -GlyGlySerPro 424  
 QY 206 TTCTTCCGACAGCCTCA---AGCCCGGCACAGACTCTGCAGCTCCAGCTCCTCCTG 150  
 Db 424 roserProSerIleSerProSerProProlIethValProSerProProthTrhProS 444  
 QY 149 CACCAAGGACAGACCCCTCCGAG-----AAGACTTCTCAT 114  
 Db 444 erProGlyGlySerProProSerProSerIleValProSerProProSerThTrhProS 464  
 QY 113 CACAGAGCAGAGCCCTGCGCTTCCACCGAGGACAGCTGGGCT----- 73  
 Db 464 erProGlySerProProthTrhSerProthTrhProthProthProGlySerProProSers 484  
 QY 72 -----CCTTCTCCGGGAATTCTGCCACAGCCTCTCAAACTGGGCTCTTGGA 24  
 Db 484 erProthTrhProthTrhProGlyGlySerProProSerSerProthTrhProthProG 504  
 QY 23 CTCGGGATCCCCGGGGCCGCG 1  
 Db 504 lyGlySerProProSerSerPro 511

```

RESULT 7
T50568
probable multi-domain regulatory protein [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50568
C:Redenbach, M.; Kleeser, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopf
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb s
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: T50568
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1334 <RED>
A:Cross-references: UNIPROT:Q9RKR9; UNIPARC:UP100000DB642; EMBL:AL133220; P1DN:GAB61705
A:Experimental source: strain A3 (2)
A:Genetics:
A>Note: SCC75A.05c

Alignment Scores:
Pred. No.: 0.0106 Length: 1334
Score: 129..00 Matches: 49
Percent Similarity: 39.0% Conservative: 13
Best local Similarity: 30.8% Mismatches: 59
Query Match: 14.0% Indels: 38
DB: Gaps: 2 7

US-10-071-510A-16 (1-493) x T50568 (1-1334)
QY 462 TCCTCAGACGAGAGCTGTGCCACAGAGGAGGCGCTGTGC-----TGAAGACCC 409
Db 300 AAlaserglyProAlaValAlaProglyserglyglyProAlaProglyTTPtrPro 319
408 CTACCAACCTTTGACTCTGGGCGCTGGGCCCTTCTCCACAGCGCCTGTAGACAGACTCA 349
Db 320 AAlaProglyThrAlaProglyserThrAlaProProHisAsp-----ThrIAsp 337
QY 348 CGCGTCAGCATGCGCTGGGAATAGGATCCATGGGATGATGAGAAATCTTGACTTTGGGA 289

```

```

Db      338  AlaAlaAspThrAla-----342
Qy      288  TGTGTTGGTGAAGAACCATTTTCTTCCCGGAATCCATCTGCATCTCGAGATGCCAGT 229
Db      343  -----ProAlaProGlyProThrSerAlaProGlyThrAlaProAlaAla 357
Qy      228  TTCGTATGAGGCTCAGACAGACTTTCTTCAGACGCTTCAGAGCCCGCCACAGACTTCCCA 169
Db      358  -----GlyThr:::-AlaAlaProAlaProGlyThrAlaGlyPro 369
Qy      168  GCTCCCTGAGCTCTCTCTGACACAGCAGCAGCACCCTCCGGAGAGACTTCTCCATCACA 109
Db      370  AlaProGlyThrSeryThrAlaProGlyThrAlaProAlaAlaGlyThrThrProAlaPro 389
Qy      108  GCCAGCCCTGCCCTTCACACAGGAGCAGCTGGGCTCT-----TCTCCG 64
Db      390  GlyThr--AlaProAlaProGlyThrAlaGlyProAlaArgAlaThrSeryThrAlaPro 408
Qy      63  GGAATTCGACACGAGCCTCMAACCTCGGGCTCTGGGACATCCGCCGGAGC 7
Db      409  GlyThrAlaProValAlaGlyThrThrProAlaProGlyThrAlaProAlaProGly 427

```

```

RESULT 8
S57447
HPBRII-7 protein - human
N/Alternate names: HPBRII-4 protein
C/Species: Homo sapiens (man)
C/Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Oct-2004
C/Accession: S57447; S57489
R/Fleischauer, K. L.
submitted to the EMBL Data Library, June 1992
A/Reference number: S57447
A/Accession: S57447
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-551 <FL>
A/Cross-references: UNIPROT:Q16630; UNIPARC:UPI000006D566; EMBL:X671336; NID:g871300; PIDN:g871300; A/Accession: S57489
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-551 <FL>
A/Cross-references: UNIPARC:UPI000006D566; EMBL:X67337; NID:g871298; PIDN:CMA47752.1; PIR:
A/Introns: 231/3
F/82-151/Domains: ribonucleoprotein repeat homology <RRM4>

Alignment Scores:
Pred. No.: 0.0145 length: 551
Score: 127.00 Matches: 67
Percent Similarity: 35.9% Conservative: 7
Best Local Similarity: 32.5% Mismatches: 77
Query Match: 13.8% Indels: 56
DB: Gaps: 10

US-10-071-S10A-16 (1-493) x S57447 (1-551)
QY 467 CCAAGTCTCCAGACGAGCATGTGGCCCCAGACAGGA----- 429
Db 189 ProGIgLyseSeSerArgIlaIaIaPheProGIngIyGIyArGIyArGIyArGPhePro 208
QY 428 GGGGCGTGtGTGGAGAGCCCTTAACCACCTCTTGACCTGGGCGCTGGC----- 381
Db 209 GIyaIaVaIaProGIyGIyASpaRgPheProGIyProIaGIyProGIyGIyProProPro 228
QY 380 CCCTTCGCCACAGCCGCTAGACAGACTCAAGCGCTCGACGATGCCTGGGA--ATAGG 324
Db 229 ProPe-----ProIaGIyIntnr-ProProAbgPrProLeuGIyProProGI 245
QY 323 ATCCATGGAGTTGATGAGAATAATCTGCATTTGGGATGTGTTGGTAAACAATTTCCT 264
Db 245 yProProGIyProProGIyProProGIyProProGIyValIeUPro-ProProLeuIaG 265
QY 263 CCCCG-----AATCACCTTCATCTCTGCGAGATGCCAGTTTCTGATGAG 219

```

[illegible]

```

Ddb 173 -----LysProProValThrLysProProThrHisT 183
Qy 173 TGCCAGCTCCCTGAGCTCTCTGACCAACGACACCTCCGAGAAAGACT---TCTC 117
Db 183 hrProSerPro---LysProProThrseryLysProThrProValThyThrProSerP 202
Qy 116 CATACACCAACGACCTGCGCTTCCACCAAGGACAGCTGGCGCTCTTCTCCGGGAATTC 57
Db 202 rOLysProProLysProSerProProThrThyThrProThrProLysProProAlaThrL 222
Qy 56 TGGCACCAACGCTCT---CAACTCGAGCTCTTGGAGACTCGGACATCCCCCG 10
Db 222 ySPProThrThrseryThrProThrHisProThrsProThrProHisThrPro 238

RESULT 11
T42731
atrophin-1 related proteogin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42731
R:Khan, F.A.; Marxolls, R.L.; Loew, S.L.; Sharp, A.H.; Li, S.H.; Ross, C.A.
submitted to the EMBL Data Library, December 1995
A:Description: CDNA sequence and expression of an atrophin-1 (DRPLA disease gene) related
A:Reference number: Z22250
A:Accession: T42731
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1006 <XUA>
A:Cross-references: UNIPROT:Q62901, UNIPARC:UPI00000E86F1, EMBL:U44091, NID:g1297310, PIR:
C:Genetics: ARP
:Gene: ARP

```

Db HisLeuSerGlyProSerProPheSerMetAlaAlaHisLeuProProProAlaLeu 407

Qy 93 CCACCA-----GGGACAGCTGGGCGCTCTCTT-----CCGGGAATT 58  
|||||  
|||||

Db 408 LysProLeuSerSerLeuSerThrHisIAsProProSerAlaHisProProLeuGln 427

Qy 57 CTGCACCAAGCCTCTCAACTCGGCGCTTTGGGACTCGGCATCCCGGGGC 7  
|||||  
|||||

Db 428 LeuMetProGlnSerGlnProLeuProSerSerProAlaGlnProGly 444

RESULT 12

H96711

hypothetical protein F14K14.17 [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

R/Theologian: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizart, L.

Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; Ker, M.; Wu, D.; Yi, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A/File: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A/Reference number: A66141; MUID:21016719; PMID:11130712

A/Accession: H96711

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-222 <STO>

A/Cross-references: UNIPROT:Q9S740; UNIPARC:UPI000009C950; GB:AE005173; NID:g6524187; PIR:

[illegible]

```

A:Gene: F1AK14.17
A:Map position: 1

Alignment Scores:
Pred. No.: 0.0423 Length: 222
Score: 121.00 Matches: 42
Percent Similarity: 33.6% Conserves: 8
Best Local Similarity: 28.2% Mismatches: 49
Query Match: 13.1% Indels: 50
DB: 2 Gaps: 5

US-10-071-510A-16 (1-493) x H96711 (1-222)

QY 411 CCCCTACCCAGCTCTTGAGCTGGGCGCTTCTCCACAGCCCTGTAACAGAC 352
    |||::: ||: ||| |||::: ||| |||
Db 60 ProValSerAlaAlaGlnProProAlaSerProValThProProAlaValThPro 79
    351 TCA-----CGGTCGACGATGCGCTGGGAATAGATCATGGAATTGATGA 307
    ::: ||| ::|||
Db 80 ThrSerProProAlaProLysVal----- 87
    306 GAAATCTGACTTGGGATGTTGTTGGTGAACAATTTCTTCCCGAATCACTTCCA 247
    ||| ||| ||| ||| |||
Db 88 -----AlaProValIleSerProAlaThnProPro 97
    246 TCCTCTGAGATGCAGATTCTTGATGAGGCTCAGCAGACTTCTTCCACAGCCTCAAG 187
    ||| ||| ||| ||| |||
Db 98 ProGlnProProGlnSer-----ProProAlaSerAla 108
    186 CCGCGCAGACTCTGCGCAGCTCCCTGAGCTCTCTCTGCACACAGGACACCTCCGGAG 127
    ||| ::| ||| ||| ||| ||| ||| |||
Db 109 ProThrValSerProProProValSerProProProAlaProThrSerProProProThr 128
    126 AAGACTTTCATCAACCGACGACCTCGGCTTCCACGAGGACACTGGGCT----- 73
    ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 ProAlaSerProProProAlaProAlaSerProProProAlaProAlaSerProProPro 148
    72 -----CCTTCTCCGGGAATTCGACCAACGCTCT 43
    ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 AlaProValSerProProProValAlaGlnAlaProSerProIleSerIleuProProAlaPro 168

```

QY 42 CAAACTCGGCTCTTGGAGTCGCGCAT 16  
Db 169 AlapProAlaPro-----Thryshis 175

RESULT 13  
C96623  
hypothetical protein F23H11.22 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: C96623  
R/Thelogs: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
anzen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maitl, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MID:21016719; PMID:11130712  
A/Accession: C96623  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-929 <STO>  
A/Cross-references: UNIPROT:Q9XIE0; UNIPARC:UPI00000A7E6A; GB:AE05173; MID:G5080823; PI  
A/Genes: F23H11.22  
A/Map position: 1

Alignment Scores:  
Pred. No.: 0.047 Length: 929  
Score: 121.00 Matches: 64  
Percent Similarity: 34.9% Conservative: 16  
Best Local Similarity: 27.9% Mismatches: 63  
Query Match: 13.1% Indels: 86  
Gaps: 13

US-10-071-510A-16 (1-493) x C96623 (1-929)

QY 473 ACCTGGCCAAAGTCTTCAGACAGAGCATGTTGCCAGACAGAGGCGCTGCTGCA 414  
Db 227 SerPheProSerSerProSerGlnIle-----HisserGlyGlyArgSer--- 242

QY 413 GACCCCTACCCACTCTTGAACCTGGGCGCTGGCCCTTCCACACCCCTGCTAGACAG 354  
Db 243 -----ProProLeuProProGlyGlnPheThrAlaGlyAsnAla---Ser 258

QY 353 ACTCAGCGTCGACGATGCTCT-----GGGATAGGATCATGCG 315  
Db 259 PheProSerSerThrGlnProProProGlyGlnTyrMetAlaGlyAsnAlaSerPhePr 278

QY 314 ATTGATGAGAATCTGACTTTGGATGTTGGTG-----AAACCATTTTCTTCC 261  
Db 278 oSerSerThrProProProGlyGlnTyrMetAlaGlyAsnAla-Probhesers 298

QY 260 CGAATCAGCTTCCA----- 247  
Db 298 erThrProLeuProProGlyGlnTyrProAlaValaAlaGlnLeuSerThSerAlaP 318

QY 246 -----TCCTCGAGATGCGAGTTTCTGATGAGGCT 216  
Db 318 roSerValProLeuProProGlyGlnTyrThrAlaValaAlaProPheSerThSer 338

QY 215 CA-----GCAGCTTCTTCCA-----GCAGCT 192  
Db 338 hGlnProValSerLeuProProGlyGlnTyrMetProGlyAsnAlaAlaLeuSerAla 358

QY 191 CAAGCCCGCC-----ACGACTCTGCGCAGCTCCGCTGCTCTCTCT----- 151  
Db 358 erThrProLeuThrProGlyGlnPheThrThrAlaAsnAlaProProAlaProProGlyP 378

QY 150 -----GCACACGCGCAGC 138  
Db 378 roAlaAsnGlnThrSerProProProProProProSerAlaAlaAlaProProPro 398

QY 137 ACCCTCCGAGAGACTTCTCCATCAGCAGCCCTGCGCTTCCACCGAGGAGCA----- 82  
Db 398 roProProProGlyGlnGlyProAlaAlaProProProProProProGlyGlyG 418

QY 81 --GCTGGCGCTCTTCTCCGGAATTCGCCACGCGCTCTCAAACTCGGCTCTTGGGA 24  
Db 418 lYAlaGlyProProProPro-----ProPheMetSerLys-----GlyP 432

QY 23 CTCGGCATCCCGCGCGCGCGCG 1  
Db 432 roProLysProProGlyAsnPro 439

RESULT 14  
T51947  
probable transcription factor HUA2 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 09-Jul-2004  
C/Accession: T51947  
R/Chen, X.; Meyerowitz, E.M.  
Mol. Cell 3, 349-360, 1999  
A/Title: HUA1 and HUA2 are two members of the floral homeotic AGAMOUS pathway.  
A/Reference number: Z25882; MID:99214850; PMID:10198637  
A/Accession: T51947  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-1392 <CHE>  
A/Cross-references: UNIPROT:Q9XER9; UNIPARC:UPI00000A0734; EMBL:AF116556; PIDN:AAD31171.1  
A/Experimental source: cultivar Columbia  
A/Genes: HUA2  
A/Map position: 5

Alignment Scores:  
Pred. No.: 0.0532 Length: 1392  
Score: 120.50 Matches: 51  
Percent Similarity: 38.5% Conservative: 14  
Best Local Similarity: 30.2% Mismatches: 63  
Query Match: 13.1% Indels: 42  
Gaps: 10

US-10-071-510A-16 (1-493) x T51947 (1-1392)

QY 386 CCGTGGCCCTTC-----TCCACAGAC-----CCTGCT 360  
Db 955 ProGlyPhePheSerSerThsAsnPheGluAspAspGluAspAspLeuProThr 974

QY 359 AGACAGACTCAGCGCTGACGATGCTGGGAATAGATCCATGATGAGAAATCC 300  
Db 975 SerGlnLysGlnLysSerThrSerAlaGlyGluArgValSerAlaLeuAspLeuGlu 994

QY 299 TGACTTTGGGATGTTGGTGAAC-----CATTTCTTCCCGGATTCATCCATCCT 243  
Db 995 ---IleHisAspThrSerSerAspLysCyHisArgValLeuGluAspValaAspH 1013

QY 242 CTGCAGATG---CGAGTTCT-----GATGAGGCTCAGCAGACTTCTT----- 202  
Db 1014 LeuGluMetGluAspValSerGlyGlnArgLysAspValaAlaProSerSerPheCyGlu 1033

QY 201 -----CCAGACGCTTCAGGCGCGCGCAG 178  
Db 1034 AsnLysThrLysGlnLysLeuAspValMetGluProValAlaGluLysSerThGlu 1053

QY 177 ACTCTGCGCACTCCC---TGAGCTCTCTCCGACCAAGGAGAGACCTCGGAG----- 127  
Db 1054 PheAsnProLeuProGluAspSerProProLeuProGlnLysSerProProProLeuPro 1073

QY 126 -----AAGACTCTCATCAGCAGCGCTGCGCTTCCACAGGAGCAGCTGGGCT 73  
Db 1074 ProLeuProProSerProProProProSerProProLeuProProProSerSerLeuProPro 1093

```
QY 72 CCTTCGCGGGAATTCTG-----CCACGACCTCTCAACTCGGCTCTT 28
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1094 ProProProlAlaAlaLeuPheProProlProProSerGlnProProPro 1113
QY 27 GGGACTCGGCATCCCCGGGCGCGCG 1
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1114 ProLeuSerProProProSerProPro 1122

RESULT 15
F86387
Probable Pto Kinase Interactor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86387
R:Thellogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hultzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzborg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86387
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-760 <STO>
A:Cross-references: UNIPROT:Q9C660; UNIPARC:UPI00000A1DD; GB:AE005172; NID:g11079512; F
C:Genetics:
A:Map position: 1

Alignment Scores:
Pred. No.: 0.0559 Length: 760
Score: 120.00 Matches: 35
Percent Similarity: 44.2% Conservative: 7
Best Local Similarity: 36.8% Mismatches: 35
Query Match: 13.0% Indels: 18
DB: 2 Gaps: 6

US-10-071-510A-16 (1-493) x F86387 (1-760)
QY 273 CCATTTTCTTCCCGAATTCATCTTCAGATGCCAGTTCTGATGAGGCTCA 214
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 ProLeuSerSerProProProGluProSerProProSerProSerLeu----- 82
QY 213 GCAGACTTCTTCACAGAGCTCAAGCCGCGCAGACTCTGCGAGCTCCTGAGCTCCT 154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 -----ThrGlyProProProThrThrLeProValSer---ProPro 95
QY 153 CCTGCACACGCGAGACCTCTCGGAGAACT--TTCATCACCAGCCAGCCT--- 100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 ProGluProSerProProProProLeuProThrGluAlaProProProAlaenProVal 115
QY 99 GCGCTTCCACACGAGGACAGCTGGGCTCTCTTCGCGGAATTCGCCACCA-----GCC 46
      ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 SerSerProProProGluSerSerProProPro-----ProProThrGluAla 132
QY 45 TCTCAAACTCGGCTCTTGAGGATCGGCAATCCCGGCGCGCGCG 1
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 ProProThrThrProThrProThrSerProSerProProThrAenPro 147
```

Search completed: April 29, 2006, 03:48:50  
Job time : 40 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 29, 2006, 03:37:25 ; Search time 33.9 Seconds  
(without alignments)  
2052.068 Million cell updates/sec

Title: US-10-071-510A-16

Perfect score: 923

Sequence: 1 cggccgcgcgcggggagtcgc.....cctgtcctcatlttgagcctgcg 493

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p model -DRV=xlp  
-Q=absbf/ABSSWEB.spool/US10071510/runat.28042006.170847.18707/app.query.fasta\_1  
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MIMATCH=0.1 -LOOFCU=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pcct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pro -NORW=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs802p  
-USER=US10071510 @CCN 1.1 466 @runat.28042006.170847.18707 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DEBLOCK=100 -LONGLOC -DEV TIMOUT=120  
-WARN TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: UniProt 05.80.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	592	64.1	489	2	Q8N7G8_HUMAN
2	586	63.5	904	2	Q6ZM23_HUMAN
3	404	43.8	895	2	Q8C117_MOUSE
4	404	43.8	975	2	Q4FZC9_MOUSE
5	305	33.0	262	2	Q8BMM1_MOUSE
6	149	16.1	555	1	GPI_CHLRE
7	144.5	15.7	911	2	Q4SM77_TETNG
8	144.5	15.7	4027	2	Q512R0_CHLIN
9	141	15.3	1143	2	Q948Y6_VOLCA
10	140.5	15.2	3889	2	Q6SS88_CHLRE
11	139.5	15.1	3409	2	Q6SS86_CHLRE
12	139	15.1	2371	2	Q58N85_CHLIN
13	138	15.0	636	2	Q7SFI5_MBUCC
14	138	15.0	676	2	Q95UC9_PIG
15	136.5	14.8	745	2	Q89X06_BRAUA
16	136	14.7	236	2	Q6QJ26_ARATH

17	136	14.7	342	2	Q6ZD62_ORYSA
18	136	14.7	731	2	Q65530_ARATH
19	135	14.6	597	2	Q5EFS5_9BRUX
20	135	14.6	659	2	Q6C708_YARLI
21	134.5	14.6	417	2	Q5ZPC7_9DELT
22	134.5	14.6	464	2	Q41645_VOLCA
23	134.5	14.6	566	2	Q95UD1_PIG
24	134.5	14.6	1701	2	Q61DV2_CABBR
25	134	14.5	511	2	Q95UD0_PIG
26	134	14.5	1495	2	Q5K1Y9_CRYNE
27	134	14.5	1525	2	Q5STB9_CRYNE
28	133.5	14.5	309	2	Q18751_CABEL
29	133.5	14.5	468	2	Q4R334_TETNG
30	133.5	14.5	687	2	Q948Y7_VOLCA
31	133.5	14.5	1997	2	Q81RM7_CHLRE
32	133	14.4	400	2	Q51HD8_MAGGR
33	132.5	14.4	698	2	Q9ASK4_ORYSA
34	132	14.3	303	2	Q5S544_CRYNE
35	132	14.3	611	2	Q412A2_GIBZE
36	131.5	14.2	326	2	Q22514_USPBD
37	131.5	14.2	763	2	Q9XDH2_MYCTU
38	131.5	14.2	1033	1	IF2_STRCO
39	131	14.2	297	2	Q527W5_MAGGR
40	131	14.2	498	2	Q5KJ55_CRYNE
41	131	14.2	3204	2	Q6X248_GALPH
42	130.5	14.1	708	2	Q9SX31_ARATH
43	130.5	14.1	1486	2	Q4VXV7_9DELT
44	130	14.1	191	2	Q82327_ARATH
45	130	14.1	191	2	Q9C550_ARATH

## ALIGNMENTS

RESULT 1  
Q8N7G8\_HUMAN  
ID Q8N7G8\_HUMAN PRELIMINARY; PRT; 489 AA.  
AC Q8N7G8;  
DT 01-OCT-2002 (TEMBLrel. 22, Created)  
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)  
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)  
DE Hypothetical protein FLJ25605.  
GN Name=Claor49;  
GN Homo sapiens (human).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;  
OC Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Thyroid;  
RA Nishimura K., Megatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Suzuki Y., Hata H.,  
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,  
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,  
RA Nagai K., Isogai T., Sugano S.,  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK098471; BAC05312.1; -; mRNA.  
DR Ensembl; ENSG00000176438; Homo sapiens.  
DR HGNC; HGNC:19861; CLorf49.  
DR InterPro; IPR002017; Spectrin.  
DR SMART; SM00150; SPEC; 1.  
DR SEQUENCE 489 AA; 56176 MW; 19ABD8DE0AC6EEF CRC64;

## Alignment Scores:

Pred. No.: 4,14e-38  
Score: 592.00  
Length: 489  
Matches: 117  
Percent Similarity: 95.2%  
Conservative: 1  
Best Local Similarity: 94.4%  
Mismatch: 6  
Query Match: 64.1%  
Indels: 0  
Gaps: 0  
DB: 2

US-10-071-510A-16 (1-493) x Q8N7G8\_HUMAN (1-489)

```
QY 3 GCGGCGCCGGGGATGCGCAGATCCCAAGAGCCGAGTTTGAGAGCGTGGTGCAGAAATTC 62
DB 192 AAGGlyProGlyAspAlaGluSerGlnGluAlaGluPheGluArgLeuValAlaGluPhe 211
QY 63 CCGGAGAGAGAGAGCCGAGCTGTCTCTGTGGAAGCCGAGGCGTGGCTGTGATGAGAAG 122
DB 212 ProGlnGluGluAlaGlnLeuSerLeuValGluAlaGlnGlyTyrLeuValMetGluYrs 231
QY 123 TCTTCTCCGAGAGGCTGCTGCGCTGTGTCAGAGAGAGCTCAGGAGCTGGCAGAGTCTGG 182
DB 232 SerSerProGlnGlyAlaAlaValAlaGlnGluGluArgGluLeuAlaGluSerTyr 251
QY 183 CCGGCTTGAAGCTGTGGAAGAAAGTCTGAGCTCATCGAAGCTGGCATCTGCAG 242
DB 252 ArgAlaLeuArgLeuLeuGluGluSerLeuLeuSerLeuLeuArgAsnTyrHisLeuGln 271
QY 243 AGGATGGAAGTGAATTCGGGGAAGAAATGCTTTTCCACCAACAATCCCAAGTCAGGA 302
DB 272 ArgMetGluValAlaAspSerGlyTyrGlyMetValPheThrAsnHisnLeuProlysserGly 291
QY 303 TTTTCATCAATCCATGATGATCTTATCCAGGATCTGTGAGCGCTGAGTCTGTACG 362
DB 292 PheLeuIleAsnProMetAspProIleProArgHisArgArgAlaAsnLeuLeuGln 311
QY 363 AGGCGTGTGGA 374
DB 312 GlnGlnGlnGly 315

RESULT 2
Q6ZM23_HUMAN PRELIMINARY; PRT; 904 AA.
ID Q6ZM23_HUMAN
AC Q6ZM23
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ16564.
GN Name=C14orf49
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Synovial membrane tissue;
RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakami B., Suzuki Y., Sugano S., Nagahara K.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahara K.,
RA Masuho Y., Nagai K., Isogai T.,
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK131436; BAD18582.1; -; mRNA.
DR HGNC:19661; C14orf49.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00435; Spectrin; 2.
DR SMART: SM00150; Spectrin; 2.
SQ SEQUENCE 904 AA; 104056 MW; CAAC7405B17CAC99 CRC64;

Alignment Scores:
Pred. No.: 1,34e-37 Length: 904
Score: 586.00 Matches: 115
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.5% Indels: 0
DB: 2 Gaps: 0

US-10-071-510A-16 (1-493) x Q6ZM23_HUMAN (1-904)
QY 3 GCGGCGCCGGGGATGCGCAGATCCCAAGAGCCGAGTTTGAGAGCGTGGTGCAGAAATTC 62
|||||
```

```
DB 612 AAGlyProGlyAspAlaGluSerGlnGluAlaGluPheGluArgLeuValAlaGluPhe 631
QY 63 CCGGAGAGAGAGAGCCGAGCTGTCTCTGTGGAAGCCGAGGCGTGGCTGTGATGAGAAG 122
DB 632 ProGlnGluGluAlaGlnLeuSerLeuValGluAlaGlnGlyTyrLeuValMetGluYrs 651
QY 123 TCTTCTCCGAGAGGCTGCTGCGCTGTGTCAGAGAGAGCTCAGGAGCTGGCAGAGTCTGG 182
DB 652 SerSerProGlnGlyAlaAlaValAlaGlnGluGluArgGluLeuAlaGluSerTyr 671
QY 183 CCGGCTTGAAGCTGTGGAAGAAAGTCTGAGCTCATCGAAGCTGGCATCTGCAG 242
DB 672 ArgAlaLeuArgLeuLeuGluGluSerLeuLeuSerLeuLeuArgAsnTyrHisLeuGln 691
QY 243 AGGATGGAAGTGAATTCGGGGAAGAAATGCTTTTCCACCAACAATCCCAAGTCAGGA 302
DB 692 ArgMetGluValAlaAspSerGlyTyrGlyMetValPheThrAsnHisnLeuProlysserGly 711
QY 303 TTTTCATCAATCCATGATGATCTTATCCAGGATCTGTGAGCGCTGAGTCTGTACG 347
DB 712 PheLeuIleAsnProMetAspProIleProArgHisArgArgArg 726

RESULT 3
Q8C117_MOUSE PRELIMINARY; PRT; 895 AA.
ID Q8C117_MOUSE
AC Q8C117
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
DE library, clone:4831426119 product:hypothetical Spectrin repeat
DE containing protein, full insert sequence.
GN Name=4831426119Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Queckenbush J.,
RA Schriml L.M., Steinhilber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barish G.,
RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaio M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombere P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohntekki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Head;
RA The FANTOM Consortium,
```



the RIKEN Genome Exploration Research Group Phase I & II Team;  
 "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 Nature 420:563-573 (2002).  
 (4)  
 NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carinci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes.";  
 Genome Res. 10:1617-1630(2000).  
 (5)  
 NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=20530913; PubMed=11076661; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carinci P.,  
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
 RA Sumi N., Iehi Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer.";  
 Genome Res. 10:1757-1771(2000).  
 (6)  
 NUCLEOTIDE SEQUENCE.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RA Adachi J., Aizawa K., Hanagaki T., Hara A., Hashizume W.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hiraoka T., Hirozane T.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Kasukawa T.,  
 RA Hori F., Imotani K., Iehi Y., Itoh M., Kagawa I., Kasukawa T.,  
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata H., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ono M., Ohnishi N., Okazaki Y.,  
 RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shingaki S., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK029216; BAC26351.1; -; mRNA.  
 DR Ensembl: ENSMUSG0000054150; Mus musculus.  
 DR MGI: MGI:2442408; 483142619gk.  
 DR InterPro: IPR002017; Spectrin.  
 DR Pfam: PF00435; Spectrin; 2.  
 DR SMART: SM00150; SPC; 2.  
 DR Hypothetical protein.  
 KW SEQUENCE 895 AA; 101521 MW; BF3E42537DA0C17 CRC64;

Alignment Scores:  
 Pred. No.: 3,13e-23 Length: 895  
 Score: 404.00 Matches: 90  
 Percent Similarity: 71.1% Conservative: 23  
 Best Local Similarity: 56.6% Mismatches: 36  
 Query Match: 43.8% Indels: 10  
 DB: 2 Gaps: 4

US-10-071-510a-16 (1-493) x Q8C117\_MOUSE (1-895)

QY 15 GATGCCAGTCCCAAGAGCGGAGTTTGAAGAGCTGTGGCAGAAATTCGCGAGAAAGAG 74  
 DB 594 AepHh1e1u8erGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 613  
 QY 75 GCCCAGCTGTCTCTGTGTGAGGCGCAGGCTGCTGTGTATGAGGAAGTCTTCTCCGAG 134  
 DB 614 ValGlnValSerLeuLeuGlnAlaLeuGlnGlnLeuValMetLeuValSerSerProGlu 633  
 QY 135 GGTGCTCCGCTGTGTGCGAGGAGGAGCTGAGGAGCTGCGAGAGCTGTGCGGAGCTTGAAG 194  
 DB 634 G1yAlaThrMetValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 653

QY 195 CTGCTGGAAGAAAGTCTGTGAGCCTCATCAGAACTGGCATCTGCAGAGTGAAGT 254  
 DB 654 LeuLeuGlnGlnValSerMetLeuSerLeuMetArgAsnGlnGlnGlnGlnGlnGln 673  
 QY 255 GATTCGGGAGAAAGATGTTTTCACCAACATCCCAAGTCCAGATTTCTCATCAAT 314  
 DB 674 AepHh1e1u8erGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 693  
 QY 315 CCCATGATCTTATTCCTCCAGGAGCATCTGCAGCCGAGTCTCTGTCAGAGGCTGTG 371  
 DB 694 ProGlnAspProIleProIleProIleProIleProIleProIleProIleProIleProIle 713  
 QY 372 GGAGAAGGGCCAGGCCCCAGGCTCAAGAGG-----TGGGTAGGGCTCTCCAGC 419  
 DB 714 ProGlnAsp--HisProGlnLeuLeuArgAspPheGlnGlnGlnGlnGlnGlnGlnGln 732  
 QY 420 ACAGGC-----CCCTCCCTGCTGTGGGCAACATGCTGCTGAGAGT 464  
 DB 733 SerIyLeuArgArgIleIleThrMetArgValAlaThrAlaIyAspLeuArgThr 751

RESULT 4  
 ID Q4FZC9\_MOUSE PRELIMINARY; PRT; 975 AA.  
 AC Q4FZC9\_MOUSE  
 DT 13-SEP-2005 (TEMBLrel. 31, Created)  
 DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)  
 DE Hypothetical protein.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Muridae; Murinae; Mus.  
 NCBI Taxid=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Eye.  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.F., Bhat N.K.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulvaney S.J.,  
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Rahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smillius D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 (2)  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Eye;  
 RG NIH MGC Project;  
 RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC099694; AA099694.1; -; mRNA.  
 KW Hypothetical protein.  
 SO SEQUENCE 975 AA; 112035 MW; 26D80A295CE0CF8B CRC64;

Alignment Scores:  
 Pred. No.: 3,17e-23 Length: 975  
 Score: 404.00 Matches: 90  
 Percent Similarity: 71.1% Conservative: 23  
 Best Local Similarity: 56.6% Mismatches: 36  
 Query Match: 43.8% Indels: 10  
 DB: 2 Gaps: 4



```

Db      1 MetcyllysserserProGlnGlyAlaThmMetValGlnGlnGlnGlySleuMet 20
QY      174 GAGTCGTGGGGGGGCTTGAGGCTGCTGGAAGAAAGGCTGCTGAGCCCTCATCAGAACTGG 233
Db      21 GlnserTrpGlnAlaLeuArgLeuLeuGlnGlnbMetLeuSerLeuMetCArgSngln 40
QY      234 CATTCGACAGAGGATGAGATGGGATTCGGGGAGAAAAAGTGTTCACCAACAACTCCCA 293
Db      41 GlnleuGlnArgThrGlnValAspThrGlyLysIysGlnValPethrAsnAsnIlePro 60
QY      294 AAGTCAGAGATTTTCATCATATCCATGGATTCCTATTCCTCCAGGCACTGTCGACGGCTGAGT 353
Db      61 LysAlaGlyPheLeuIleAsnProGlnAspProIleProArgArgGlnAlaGlyAlaAsn 80
QY      354 ---CTGCTAGACAGAGGCTGCGAGAGAGGGGCCAGGCCCAAGCTCAGAGAG----- 401
Db      81 ProleuGlnGlyHisAspLeuProGlnAsp---HisProGlnLeuLeuArgAspPheGlu 99
QY      402 ---TGGTAGGGGCTCTCCAGACACAGGC-----CCCTCCCTGCTCGGGGACACA 446
Db      100 GlnTrpLeuGlnAlaGlnAspSerLysLeuArgArgIleIleThrMetArgValAlaThr 119
QY      447 TGCTCTGCTCGAGAGCT 464
Db      120 AlaLysAspLeuArgThr 125

RESULT 6
GPI_CHURE
ID      GPI_CHURE      STANDARD;      PRT;      555 AA.
AC      O9PFO6; Q03927;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      10-MAY-2005 (Rel. 47, Last annotation update)
DE      Vegetative cell wall protein gpi precursor (Hydroxyproline-rich
DE      glycoprotein 1).
GN      Name=GPI;
OS      Chlamydomonas reinhardtii.
OS      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC      Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
NCBI_TaxID=3055;
[1]
RN      NUCLEOTIDE SEQUENCE.
RP      MEDLINE=21159092; PubMed=11258910; DOI=10.1021/bi0023605;
RX      Ferris P.J., Woessner J.P., Waffenschmidt S., Kilz S., Drees J.,
RA      Goodenough U.W.;
RA      "Glycosylated polypurine II rods-with-kinks as a structural motif in
RT      plant hydroxyproline-rich glycoproteins."
RL      Biochemistry 40:2978-2987(2001).
[2]
RN      PRELIMINARY PARTIAL NUCLEOTIDE SEQUENCE.
RP      MEDLINE=91017504; PubMed=1699225;
RX      Adair W.S., Apt K.E.;
RA      "Cell wall regeneration in Chlamydomonas: accumulation of mRNAs
RT      encoding cell wall hydroxyproline-rich glycoproteins."
RL      Proc. Natl. Acad. Sci. U.S.A. 87:7355-7359(1990).
-1- FUNCTION: Major component of the outer cell wall w6 (crystalline)
layer.
-1- SUBUNIT: Associates with GP2 and GP3.
-1- PTM: N-glycosylated and O-glycosylated.
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      EMBL, AF309494; AAC45420.1; -; Genomic DNA.
DR      EMBL, M58486; AAA69706.1; ALT_SEQ; mRNA.
DR      GLOSUSuiteDB; O9PFO6; -.
DR      InterPro, IPR003882; PfamExtensin.
DR      PRINTS; PRO1218; PSTIEXTENSIN.
KW      Glycoprotein; Repeat; Signal.

```

FT SIGNAL	1	29	Potential.
FT CHAIN	30	555	Vegetative cell wall protein gpl.
FT REGION	40	339	49 X 5 AA approximate repeats of P-P-S-P-X.
FT COMPART	259	279	Poly-Pro.
FT CARBOHYD	392	399	N-linked (G1CNC. . .) (Potential).
FT CARBOHYD	435	455	N-linked (G1CNC. . .) (Potential).
FT CARBOHYD	493	493	N-linked (G1CNC. . .) (Potential).
FO SEQUENCE	555 AA;	54220 NM;	6A5849A04655052F5 CXC64;

**Alignment Scores:**

Pred. No.:	0.00397	555
Score:	149.00	Matches:
Percent Similarity:	36.8%	Conservative:
Best Local Similarity:	34.9%	Mismatches:
Query Match:	16.1%	Indels:
DB:	1	Gaps:
		6

US-10-071-510A-16 (1-493) x GP1\_CHLRE (1-555)

446 TGTGCCCCAGACAGGAGGGCC-----TGTGCTGGAGACCCCTACCCACTTGTAC 393

Db 31 CysValPro-----GlyGlyIlePheAsnCysProProSerProAlaProProSerPro

392 CTGGGGCCCTGGCCCCCTCTCCCAAGCCCTGCTAGACAGACICACGCGTCGACGATGCC 353

D6 49 AAlProProSerProAlaProProSerProAlaProIofIofSerfIofAlaIofIofccccc 00

[illegible]

214

```

75 ProProProProProPro]aProProSer-----Pro 86

```

213 GCGAGACTTTCTTCCAGCAGCCCTCAAGGCCCGCCGACGACTCTGCCAGCTCCCTGAGCTCCT 15

db

87 AlaProProSerProAlaProProSerProAlaProProSerProAlaProProSerPro 100

153 CCTGCACCA CGGAGCACCCTCCGGAGAGACTTCTCCATCACCAGCCAGCCCTGCGCTT 94

Db 107 AlaProP SerProAlaProP SerProP SerProP SerProAlaPro 12

93 CCA---CCAGGACAGCTGGGCTCTCTTCGCGGAATTCTGCCACCCAGCCTCTCAACT 37

```

Db      127 ProSerProSerProProAlaProProSerProSer-----ProProSerProAla 14

```

36 CGGCTCTTGGGACTCGGCATCCCCCGGGCCGGG 1

Db 144 PROPROLeuPROserPRIOAlaPRIOProsePRIO 155

RESULT /  
Q4SMF7 TETING

ID	QASM67; TELING ENLIGHTENMENT /	
AC	QASM67;	

DT	13-SEP-2005 (Tremblay, 31, Last sequence update)
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

DE Chromosome undetermined SCAF14551, whole genome shotgun sequence.  
CN OPENames-CSTENG00015781001;

OS Tetraodon nigroviridis (Green putter).  
OC Eukaryota: Metazoa: Chordata: Vertebrata; Euteleostomi;

0C Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
0C Actinopterygii; Neopterygii; Ieiostei; Euteleostei; Neoteleostei;  
0C Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

```

0C  Ictidionomidae; Ictidionomidae; Ictidionomidae;
0X  NCBI_TaxID=99883;

```

RP NUCLEOTIDE SEQUENCE. (2) Purified P. horridus P. horridus T. Strang-Kaplan, N. J.

RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A., Taffe D., Fisher S., Iufalla G., Dossat C., Seurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,



```

Db      1377 ProSerProGlnProProSerProSerProSerProAlaProAlaProProSerProAlaProLeuGlu 1396
Qy      51 CCAGCCTCTCAAACTCGGCGCTCTGGAGCTCGGCAATCCCGCGGCGCG 1
Db      1397 ProAlaAlaProAlaProProGlyProProProGlnProProGlyAlaPro 1413

RESULT 9
Q948Y6_VOLCA
ID      Q948Y6_VOLCA PRELIMINARY; PRT; 1143 AA.
AC      Q948Y6_
DT      01-DEC-2001 (TRMBLrel. 19, Created)
DT      01-DEC-2001 (TRMBLrel. 19, Last sequence update)
DT      01-MAR-2004 (TRMBLrel. 26, Last annotation update)
DE      VMP4 protein.
GN      Name=vmp4;
OS      Volvox carterii f. nagariensis.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC      Chlamydomonadales; Volvocaceae; Volvox.
OX      NCBI_TaxID=3068;
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=HK 10;
RA      MEDLINE=21382906; PubMed=11489172;
RA      Hallmann A., Amon P., Godl K., Heitzer M., Sumper M.;
RT      "Transcriptional activation by the sexual pheromone and wounding: a
RT      new gene family from Volvox encoding modular proteins with
RT      (hydroxy)proline-rich and metalloprotease homology domains.";
RL      Plant J. 26:583-593(2001).
EMBL    AJ311550; CAC39319.1; -; mRNA.
DR      MEROPS; M11.002;
DR      GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR      InterPro; IPR003014; PAN.
DR      InterPro; IPR003609; Pan_app.
DR      InterPro; IPR008752; Peptidase_M11.
DR      InterPro; IPR003882; PstcII_extensin.
DR      InterPro; IPR000772; Ricin_B_lectin.
DR      Pfam; PF00024; PAN; 1.
DR      Pfam; PF05548; Peptidase_M11; 1.
DR      PRINTS; PRO1218; PSTLEXTENSIN.
DR      PRINTS; PRO0652; Ricin_B_lectin; 2.
DR      PROSITE; PS50948; PAN; 2.
DR      PROSITE; PS50231; RICIN_B_LECTIN; 2.
SQ      SEQUENCE 1143 AA; 123155 MW; 0CE0FD2CBD52200C CRC64;

Alignment Scores:
Pred. No.: 0.0188 Length: 1143
Score: 141.00 Matches: 42
Percent Similarity: 32.3% Conservative: 10
Best Local Similarity: 26.1% Mismatches: 55
Query Match: 15.3% Indels: 54
Gaps: 4

US-10-071-510A-16 (1-493) x Q948Y6_VOLCA (1-1143)
Qy      483 ATGAGACAGAGCGCAATGCTCTGAGCAGACAGATGTCGCCAGAGGAGGCG 424
Db      486 MetleuclulglyllealavallyrthserProProhevalProProProProSerPro 505
Qy      423 CTGTGTGAGACCCCTTACCTCTTGAACCTGGGGCTGGCCCTTCTCCACAGCCC 364
Db      506 LeuLeuThierProArgProProSer-----ProArgProProArgPro 520
Qy      363 TGTCTAGACAGACTCAGCCGCTGAGATGCTGGGATATGATTCATGGATTGATGAA 304
Db      520 ----- 520
Qy      303 ATCTGACTTTGGAGTGTGTGTGTAACAACATTTCTTCCCGAATTCATTCATCC 244
Db      521 -----SerProProSerProProPro----- 527
Qy      243 TCTGCAATGCCAGTTTCTGTATGAGGCTCAGACAGATTTCTTCCAGACGCTCAAGGCC 184

```

```

Db      528 -----ProProSerProProPro 533
Qy      183 GCCAGACTCTGCCAGCTTCCTGAGCTCTCTGCAACCAAGCAGACACCTTCGGAGAAG 124
Db      534 ProProSerProProProProProProProProProProProProProProPro 553
Qy      123 ACTTCTCATACGACGCCAGCTTCCAGCCAGGAGACAGTGGGCTCTCTCCG 64
Db      554 SerProProProProProProProProProProProProProProProProPro 573
Qy      63 GGAATTCGACACGAGCTCTCAAACTCGGCTTGGGACTCGGCAATCCCGGCGCG 4
Db      574 Pro-----ProProProSerProProProProProProProArgHisProProSerPro 591
Qy      3 CCG 1
Db      592 Pro 592

RESULT 10
Q6SSB8_CHLRE
ID      Q6SSB8_CHLRE PRELIMINARY; PRT; 3889 AA.
AC      Q6SSB8;
DT      05-JUL-2004 (TRMBLrel. 27, Created)
DT      05-JUL-2004 (TRMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TRMBLrel. 27, Last annotation update)
DE      Minus agglutinin.
GN      Name=SAD1;
OS      Chlamydomonas reinhardtii.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC      Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX      NCBI_TaxID=3055;
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CC-621;
RA      PubMed=15659633; DOI=10.1105/tpc.104.028035;
RA      Ferris P.J., Welfenschmidt S., Umen U.G., Lin H., Lee J.H., Ishida K.,
RT      "Plus and Minus Sexual Agglutinins from Chlamydomonas reinhardtii.";
RL      Plant Cell 17:597-615(2005).
DR      EMBL; AY450929; AAS07042.1; -; Genomic DNA.
DR      GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR      InterPro; IPR002951; Atrophin.
DR      InterPro; IPR006315; Autotransporter.
DR      InterPro; IPR003882; PstcII_extensin.
DR      PRINTS; PRO1222; ATROPHIN.
DR      PRINTS; PRO1218; PSTLEXTENSIN.
DR      TIGRFAMs; TIGR01414; autotrans_bar1; 1.
SQ      SEQUENCE 3889 AA; 389223 MW; 4DE5B44D5507214A CRC64;

Alignment Scores:
Pred. No.: 0.0243 Length: 3889
Score: 140.50 Matches: 43
Percent Similarity: 33.1% Conservative: 4
Best Local Similarity: 30.3% Mismatches: 42
Query Match: 15.2% Indels: 53
Gaps: 5

US-10-071-510A-16 (1-493) x Q6SSB8_CHLRE (1-3889)
Qy      411 CCCTACCCACACCTTGTGAGCTGGGCGCTTCTCTCCACAGCCCTGTAAGACAG 352
Db      1369 ProProProProProPro-----ProThProProSerPro----- 1379
Qy      351 TCACGGGTGACGAGTCTGGGAATGATTCATGGATTGATGAAATCTGACTTTG 292
Db      1379 ----- 1379
Qy      291 GGAATGTTGTTGGTAAACCATTTTCTTCCCGAATTCATTCCTCTGAGATGCC 232
Db      1380 -----GlnProProValProProSerProThProProSer----- 1391
Qy      231 AGTTTCTGATGAGGCTCAGAGACTTCTTCCAGACGCTCAAGGCCGCGCAGACTCTG 172

```

```
DB 1392 -----ProGlnProProserProAlaProProser 1401
QY 171 CCAAGTCTCTGAGCTCTCTCGACACGAGCAGACCTTCGGAGAAAGATTTCATCA 112
DB 1402 ProAlaProSerAlaProLeuGlnProSerProAspProProserProGlnProProser 1421
QY 111 CCAGCC-----AGCCCTGGGCTTCACACGAGGAGAGCTGGGCTCTCTCT 67
DB 1422 ProAlaProGlnProProProserProProProserProSerThrProProserProProser 1441
QY 66 CCGGGAATTCTGACACGAGCTCTCAAACTCGGCTCTTGGAATCGGCATCCCGCGGC 7
DB 1442 ProAlaProLeuAlaProAlaProProValProProMetAlaProGln---ProProser 1460
QY 6 CCGCCG 1
DB 1461 ProPro 1462

RESULT 11
Q6SSB6_CHLRE PRELIMINARY; PRT; 3409 AA.
ID Q6SSB6_CHLRE
AC Q6SSE6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Plus agglutinin.
GN Name=SAG1;
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxId=3055;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15659633; DOI=10.1105/tpc.104.028035;
RA Ferris P.J., Waffenschmidt S., Umen J.G., Lin H., Lee J.H., Ishida K.,
RA Kubo T., Lau J., Goodenough U.W.;
RT "Plus and Minus Sexual Agglutinins from Chlamydomonas reinhardtii.";
RT Plant Cell 17:597-615(2005).
DR EMBL; AY450930; AAS0704.1; -; Genomic DNA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR002951; Atropin.
DR InterPro; IPR003882; Pstc11_extensin.
DR PRINTS; PR01223; ATROPHIN.
DR PRINTS; PR01218; PST1EXTENSIN.
SQ SEQUENCE 3409 AA; 336045 MW; 7PE87633ED6631F CRC64;

Alignment Scores:
Pred. No.: 0.0286 Length: 3409
Score: 139.50 Matches: 53
Percent Similarity: 37.4% Conservative: 14
Best Local Similarity: 29.6% Mismatches: 57
Query Match: 15.1% Indels: 55
DB: 2 Gaps: 7

US-10-071-510A-16 (1-493) x Q6SSB6_CHLRE (1-3409)
QY 467 CCAAGTCTCTGAGCAGACGATGTTGCCACAGAGGAGGGCTGTGCTGAGACCCC 408
DB 505 ProSerProProSerProProSerProPro-----Pro 515
QY 407 TACCCACCTTGACCTGGGGCTGGCCCTTCTCCACAGCCCTGTAGACAGACTAC 348
DB 516 SerProGlnProProSerProProProLeuProProSerProProSerProThrPro 535
QY 347 GCGTCGACGATCGCTGGGATAGATCCATGGGATTGATGAAATCTGACTTTGGAT 288
DB 536 ValAla-ArgCysIleGlnValGly-----Gly11 545
QY 287 GTTGTGGTGAACCAATTTCTTCCCGAATCCACTTCATCTCT----- 241
DB 545 e-Cys-----AspSerProSerProMetProProSerProArgProProG 560
QY 240 -----GCAGATG 234
```

```
DB 560 InProProserProProProProProProArgProProProAlaProArgProSerP 580
QY 233 CCAAGTTTCTGATGAGGCTCAGACACTTTCTTCAGACGACTCAAGCCGCCAGACTC 174
DB 580 roProBehisProProSerProAspSerProProAlaSerValPro----- 596
QY 173 TGCAGCTCCCTGAGCTCTCTCGACACGAGCAGACCCCTCGGAGAAAGATTCTGCAT 114
DB 597 --ProSerPro--GlnProProSerProLys-----ProProserProAlaProPro 613
QY 113 CACACGACGAGCCCTCGCTTCACACGAGGAGACGCTGGGCTCTCTTCGAGGAATTCTGC 54
DB 613 erProAlaProProserProProProProProProProAlaProProserProAlaProPro 633
QY 53 CACAGCTCTCAAACTCGGCTCTTGGAATCGGCAATCCCGGCGCGCG 1
DB 633 erProAlaProProserProGlnProProserProProValProProGlnProPro 650

RESULT 12
Q58NA5_CHLIN PRELIMINARY; PRT; 2371 AA.
ID Q58NA5_CHLIN
AC Q58NA5;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Plus agglutinin (Fragment).
GN Name=SAG1;
OS Chlamydomonas incerta.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
OC Chlamydomonadales; Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxId=51695;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CC-1870;
RA Lee J.-H., Ferris P.J., Waffenschmidt S., Ishida K., Goodenough U.W.;
RT "Evolution of Genes Encoding Chimeric Hydroxyproline-Rich
RT Glycoproteins and Sex-Related Proteins in Two Species of
RT Chlamydomonas.";
RT Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY937239; AAX33674.1; -; Genomic DNA.
FT NON TER 2371
FT 2371
SQ SEQUENCE 2371 AA; 236526 MW; A6B5F15927378BA CRC64;

Alignment Scores:
Pred. No.: 0.0298 Length: 2371
Score: 139.00 Matches: 50
Percent Similarity: 35.4% Conservative: 6
Best Local Similarity: 31.6% Mismatches: 47
Query Match: 15.1% Indels: 56
DB: 2 Gaps: 6

US-10-071-510A-16 (1-493) x Q58NA5_CHLIN (1-2371)
QY 473 ACGTGCCCAAGCTCCAGACGATGTTGCCACAGAGGAGGGCTGTGCTGGA 414
DB 1215 ThrProProAlaProGlnProProSerProProAlaProProAla----- 1228
QY 413 GACCCCTTCCACCTTGACCTGGGGCTGGCCCTTCTCCACAGCCCTGTAGACAG 354
DB 1229 --ProGlnProProSerProAlaProProSerProAnProProSerProAlaProThr 1247
QY 353 ACTCAGCGGTGACGATGCTGGGAAATG-GATCCATGGGATGTAGGAAATCTGACT 295
DB 1248 ThrProAlaSerProGlnProProSerProGlnPro----- 1259
QY 224 TTGGATGTGTGTGTAACCAATTTCTTCCCGAATCCACTTCATCTGCAAT 235
DB 1260 -----ProSerProSerPro----- 1264
QY 234 GCGAGTTTCTGATGAGGCTCAGACACTTTCTTCAGACGAGCTCAAGCCGCCAGACT 175
DB 1265 -----ProValProProserProAlaProPro 1273
```

QY 174 CTGCCAGCTCCTGAGCTCTCTGACACGACGACCTTCGAGAGAATTCTTCCA 115  
 |||||  
 DB 1274 SerProAlaProLeuProProSerPro-----AspProPro-----SerPro 1288  
 |||||  
 QY 114 TCACCAAGCCGCTGCGCTTCCACAGGAGACGCTGGGCTCTTCTTCGGAATTCTG 55  
 |||||  
 DB 1289 ValProProSerProAlaProProProSerProProAlaProProSerProGluProIle 1308  
 |||||  
 QY 54 CCACCAAGCTCTCAAACTCGGCTCTGGGACTCGGCATCCCCGGCGCGCG 1  
 |||||  
 DB 1309 ProProAlaPro-----ProProSerProPro 1317  
 |||||

## RESULT 13

Q7SF15\_NEUCR PRELIMINARY; PRT; 636 AA.  
 ID Q7SF15;  
 AC Q7SF15;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN Name=NCU07438.1;  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=OR744;  
 RA Gallaen J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,  
 RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,  
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,  
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,  
 RA Seltrenikoff C.P., Kinsey U.A., Braun E.L., Zelter A., Schulte U.,  
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,  
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gueire S.,  
 RA Kamal M., Kamysseil M., Mauceli E., Bielek C., Rudd S., Fishman D.,  
 RA Kyrstofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,  
 RA Cognoli C., Macino G., Catchside D., Li W., Pratt R.J., Osmann S.A.,  
 RA Desouza C.C., Glass L., Otbach M.J., Berglund J., Voelker R.,  
 RA Vardona O., Plamann M., Selter S., Dunlap J., Radford A., Aramayo R.,  
 RA Natvig D.O., Alex L.A., Manhaupt G., Ebbole D.J., Fretlag M.,  
 RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,  
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";  
 RL Nature 0-0-0(2003).  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 CC -1- SUBUNIT: Isoform 1 and isoform 2 encode coiled-coil structures  
 CC that mediate homo- and heteromultimerization (By similarity).  
 CC -1- DOMAIN: The WH1 domain interacts with the PPXXF motif in GRM1,  
 CC GRM5, RFR1, RYR2, ITPR1, SHAK1 and SHAK3 (By similarity).  
 CC EMBL: AABX0100025; EAX3589.1; -; Genomic\_DNA.  
 DR InterPro: IPR000697; EVH1.  
 DR InterPro: IPR011993; PH type.  
 DR InterPro: IPR001960; WH1.  
 DR InterPro: IPR003124; WH2.  
 DR Pfam: PF00568; WH1; 1.  
 DR Pfam: PF02205; WH2; 1.  
 DR Coiled coil; Hypothetical protein.  
 DB: SEQUENCE 636 AA; 66201 MW; AE102B813CA938B6 CRC64;

## Alignment Scores:

Pred. No.: 0.0299 Length: 636  
 Score: 138.00 Matches: 39  
 Percent Similarity: 43.8% Conservative: 3  
 Best Local Similarity: 40.6% Mismatches: 36  
 Query Match: 15.0% Indels: 18  
 DB: Gaps: 4

US-10-071-510A-16 (1-493) x Q7SF15\_NEUCR (1-636)

QY 255 CCACTTCATCCTCTGAGATGACGATTCTTGATGAGGCTCAGACACTTTCTT---CCA 199

DB 407 ProLeuProProLysAlaPro-----GlyProAlaProProLeuProPro 421  
 |||||  
 QY 198 GCAGCTCAAGAGCCGCCACGACACTTCGACAGCTCTCTCTGACACGAGCAG 139  
 |||||  
 DB 422 AlaSerSerArgProProProMetLeuProThr-----ArgSerProAlaProProGln 439  
 |||||  
 QY 138 CACCTTCGAGAGAACTTCTTCATCACCAGCAGCCCTGGCGCTTCCA----- 91  
 |||||  
 DB 440 AlaProProLeuProThrSerAlaAlaProProProProLeuProAlaThrGlnAla 459  
 |||||  
 QY 90 -----CAAGGACAGCTGGGCTCTCTTCGCGGAATTCTGCACCA 49  
 |||||  
 DB 460 ProProProProProProLeuProAlaThrSerAlaProProProProProProAlaProPro 479  
 |||||  
 QY 48 GCCTCTCAAACTCGGCTCTTGGGACTCGGACTCCCGCGCGCGCG 1  
 |||||  
 DB 480 AlaProProAlaProProLeuProAlaAlaHisAlaProProProPro 495  
 |||||

## RESULT 14

Q95JC9\_PIG PRELIMINARY; PRT; 676 AA.  
 ID Q95JC9;  
 AC Q95JC9;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Basic proline-rich protein.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;  
 OC Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Parotid gland;  
 RA PubMed=15805110;  
 RX Zhang Q., Sealay A.A., Tieche J.M., Kye-yune-Nyombi E., Sands J.F.,  
 RA Oberg K.C., Leonora J.,  
 RT "Cloning and Functional Study of Porcine Parotid Hormone, a Novel  
 RT Proline-rich Protein.";  
 RL J. Biol. Chem. 280:22233-22244(2005).  
 DR EMBL: AY035849; AK61383.1; -; mRNA.  
 DR GO: GO:0016021; C:integral to membrane; IEA.  
 DR GO: GO:0005199; F:structural constituent of cell wall; IEA.  
 DR GO: GO:0005249; F:voltage-gated potassium channel activity; IEA.  
 DR GO: GO:0006813; P:potassium ion transport; IEA.  
 DR InterPro: IPR002951; Atrophin.  
 DR InterPro: IPR005404; KVs\_3 channel.  
 DR InterPro: IPR003882; PstR1L\_extensin.  
 DR PRINTS: PR01222; ATROPHIN.  
 DR PRINTS: PR01582; KV33CHANNEL.  
 DR PRINTS: PR01218; PSTLEXTENSIN.  
 DB: SEQUENCE 676 AA; 62297 MW; 3008BC41EFD81FC9 CRC64;

## Alignment Scores:

Pred. No.: 0.0301 Length: 676  
 Score: 138.00 Matches: 60  
 Percent Similarity: 36.4% Conservative: 3  
 Best Local Similarity: 34.7% Mismatches: 58  
 Query Match: 15.0% Indels: 52  
 DB: Gaps: 11

US-10-071-510A-16 (1-493) x Q95JC9\_PIG (1-676)

QY 467 CCAAGTCTCAGAGACGATGTCGCCAGACGAGGCGCTTGCTGAGAC--- 411  
 |||||  
 DB 417 ProProProProProProAlaAlaGluProGlnGlnGlyProAlaProSerGlyAlaPlys 436  
 |||||  
 QY 410 -----CCCTACCACTCTTGACTGGG-----CTGGCCCTCTTCTCCACACG 366  
 |||||  
 DB 437 ProLysLysLysProProProProAlaGlyProProProProGlyProProSerProGly 456  
 |||||  
 QY 365 CCGTCTGAGACGACTCAGCGCTGAGACATGCCCTGGAGAAATGATCCATGGGATTGATGAG 306  
 |||||

```

Db      457 ProAla-----ProProGlyAlaArg--Pro----- 464
Qy      305 AATCTGACTTTGGATGTTGTGTGTAAMACCATTTTCTCCCGAATCCACTTCAT 246
Db      465 -----ProProGlyProProProGlyProP 474
Qy      245 CCTCTGCAGATGCCATTTTCTGTATGAGGCTCAGCAGACTTCTTCCAGCAGCCTCAAGC 186
Db      474 roProPro-----GlyProAla-----ProProGlyAlaArgP 485
Qy      185 CCGCCAGACTCTCCAGCTCCC-----TGAGCTCTCTCCGACACAC 144
Db      485 roProProGlyProProProProGlyProProProProGlyProAlaProProGlyAlaA 505
Qy      143 GCGACGACCTT---CCGAGAGAGACTTTCATCACCAGCCAGCCTGCGCTTCCACAG 87
Db      505 rgProProProGlyProProProProGlyProProProProGlyProAla---ProProG 524
Qy      86 GGACA-----GCTGGGCTCTTCCCGGAATTCTGCACAGCAGCCTCTCAACATC 36
Db      524 lYAlaArgProProProGlyProProProProGly-----ProProProProGlyProA 542
Qy      35 GGCCTCTTGGAAGCTCCGACATCCCGCGGCGCGCG 1
Db      542 laProProGlyAlaArgProProProProGlyProPro 553

RESULT 15
Q89X06 BRAJA PRELIMINARY; PRT; 745 AA.
AC Q89X06;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE B1r0521 protein.
GN OrderedLocusNames=B1r0521;
OS Bradyrhizobium japonicum.
OC Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2484998; PubMed=12592725;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iizuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Ref. 9:189-197(2002).
EMBL; BA000040; BAC45786.1; -; Genomic_DNA.
DR HSSP; P0176; 1OAP.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR InterPro; IPR006665; OmpA/MocB.
DR InterPro; IPR003882; Pstc1_extensin.
DR Pfam; PF00691; OmpA; 1.
DR PRINTS; PR01218; PSTLEXTENSIN.
DR ProDom; PDO00930; OmpA/MocB; 1.
KM Complete proteome.
SQ SEQUENCE 745 AA; 74544 MW; 155EDFCC74DBC6D CRC64;

```

## Alignment Scores:

```

Pred. No.:      0.0401      Length:      745
Score:          136.50      Matches:      48
Percent Similarity: 35.5%      Conservative: 12
Best Local Similarity: 28.4%      Mismatches: 65
Query Match:     14.8%      Indels:      43
DB:              2          Gaps:          6

```

US-10-071-510A-16 (1-493) x Q89X06\_BRAJA (1-745)

```

Qy      441 CCCAGAGAGGAGGGGCGCTGTGTGAGACCCCTACCACCTTGTGAGCTGGGCGCTGG 382
Db      181 ProProAlaAGlyProAlaAlaArgProThrProAlaProThrAlaThrPro-----Thr 198
Qy      381 CCCCTTCTCCACAGCCCTGTAGACAGACTCAGCGCTGACGATGCTGGGAATAGAT 322
Db      199 ProValAlaProProProAlaAlaProThrAlaArgProGlySerPro----- 214
Qy      321 CCATGGATGTGTAAGAAATCTGACTTTGGATGTGTGTGTAAMACCATTTCTTCC 262
Db      215 -----AlaProAlaAlaThr 219
Qy      261 CCGAATCCACTTCATCTCTGCA-----GATGCCAGTTTCTGA 223
Db      220 ProAlaProThrProThrProAlaProThrAlaThrProAlaProThrAlaThrProAla 239
Qy      222 TGAGGCTCAGAGACTTCTTCCAGCAGCCTCAAGGCCGCGACG----- 178
Db      240 ProGlySerThrProGlyAlaBProProAlaGlyArgProGlyAlaBProProGlyVal 259
Qy      177 -----ACTTCCGAGCTTCCCTGAGCTCTCTCTGACACAGCGACACACCTCCGAG 127
Db      260 ArgProGlySerProProAlaAlaGlySerProProAlaProGlyAlaThrProAlaPro 279
Qy      126 AAGACTTCTCATCACACAGCCAGCCCTGCGCTTCA-----CCAGGAGACAGCT 79
Db      280 ThrThrThrProAlaProGlyGlyThrAlaThrProProSerGlyArgProGlyProAla 299
Qy      78 GGGCCTCTTCTCGGGAATTCTGCACACAGCTCTCAAACTGGCCTT----- 28
Db      300 SerThrProAlaProGlyAlaAlaThrProAlaProThrAlaThrProAlaProGlyGly 319
Qy      27 GAGACTCGGACATCCCGCGGCGCGCG 1
Db      320 AlaLeuThrProProProGlyArgPro 328

```

Search completed: April 29, 2006, 03:47:43  
 Job time : 182.5 secs





```
QY 258 AATCCATTCATCCTCTGCGATGCGAGTTCTGATGAGGTCCAGACATTTCTTCCA 199
    ||| :|||:|||||
Db 125 AaRthMePrAlaSerAla-----AlaRthGlnValAlaGhS 137
QY 198 GCAGCCTCAAGGCGCGCCGAGACTTCCAGCTCC-----TGAGCTCTCTTGA 148
    |||:|||||
Db 138 AlaAlaPrGlnPrAlaAlaAla---ProAlaPrValAlaAlaAlaPrGln 156
QY 147 CCACGGCAGACCCCTCCGAGAGAACTTCTCCATCA----- 112
    |||
Db 157 AlaArPrAlaPrObPrGlnAlaArPrGlnAlaPrObPrGlnAlaPrO 176
QY 111 -----CCAGCCAGCCCTCGGCTTCCA 91
    |||
Db 177 ArGPrObPrGlnAlaGlnValPrObPrObPrGlnAlaArPrObPrGlnAlaPrO 196
QY 90 CCAGGAGCAGCTGGGCTCTTCTCCGGAATTCTGCCAGCAGCTCT-----CAACT 37
    |||:|||||
Db 197 PrGlnMetAlaArPrObPrGlnAlaArPrObPrGlnMetPrObPrObPrGlnAlaPrO 216
QY 36 CGGCTCTTGAGACTCGGCATCCCGCGGC-----CGGCG 1
    |||
Db 217 ArGPrObPrGlnAlaPrGlnMetPrObPrGlnMetAlaArPrO 230

RESULT 2
US-09-949-016-7961
/ Sequence 7961, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7961
/ LENGTH: 553
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-7961

Alignment Scores:
Pred. No.: 0.000936 Length: 553
Score: 127.00 Matches: 67
Percent Similarity: 35.9% Conservative: 7
Best Local Similarity: 32.5% Mismatches: 77
Query Match: 13.8% Indels: 56
DB: 2 Gaps: 10

US-10-071-510A-16 (1-493) x US-09-949-016-7961 (1-553)
QY 467 CCAAGCTCCAGAGAGAGCATGTCGCCACAGACAGGA----- 429
    |||:|||||
Db 191 PrGlnGlnSerSerArGlnAlaAlaPrGlnGlnGlnAlaArGlnAlaArGlnArGlnPrOb 210
QY 428 GGGGCTGTGCTGAGACCCCTTACCACTTGAACCTGGGCGCTGAC----- 381
    |||
Db 211 GlnAlaValPrObPrGlnAlaArGlnPrObPrGlnAlaArGlnPrObPrGlnPrObPrO 230
QY 380 CCGTTTCCACAGCCCTGCTAGACAGCTCAGCGCTCGACATGCCCTGGGA---ATAG 324
    |||:|||||
Db 231 PrObPrO-----PrObAlaGlnGlnThr-PrObArGPrObPrObGlnPrObPrO 247
QY 323 ATCCATGGGATTTGATGAGAAATCCTGACTTGGGATGTTGTTGTAACCATTTTCTT 264
    |||:|||||
```

```
Db 247 YPrObPrGlnPrObPrGlnPrObPrObPrGlnValLeuPro-PrObPrOLeuAlaG 267
QY 263 CCGCG-----AATCCATTCGATCCTCGAGATGCGAGTTCGTGATGAG 219
    |||:|||||
Db 267 YPrObPrAsnArGlnAlaArGlnPrObPrO-PrObAlaLeuPhePrGlnPrObPrO 286
QY 218 GCTCAGACAGACTTCTTCAGAGAGCTCGAGCCGCGCAGC-ACCTGCCAGCT----- 166
    |||
Db 287 GlnGlnPrObPrOLeuGlnPrObLeuPrObPrGlnPrObPrObPrObPrObPrGlnPrO 306
QY 165 ---CCCTAGACTCTCTCTGACACGAGCAGACCTCCGGAGAACTTCTCCATCACA 109
    |||
Db 307 PrObPrObPrGlnPrObPrObPrGlnGlnGlnPrObPrObPrObPrGlnPrObPrO 326
QY 108 GCCACCGCT-----GGGCTTCACACCGAGGACACT 79
    |||
Db 327 PrObArGPrObPrGlnPrObLeuGlnPrObPrOLeuThrLeuAlaPrObPrObPrObPrO 346
QY 78 GGGCTCTCTCTCCGGAATTCTGCCACAGCAGCTCTCAAACTCGGCTCTT----- 28
    |||:|||||
Db 347 GlnPrObPrO-----PrGlnAlaPrObPrObPrObPrObPrObPrObPrObPrObPrO 365
QY 27 -----GGGACTCGGCATCCCGC----- 10
    |||
Db 366 PrObPrOThrAsnSerGlnMetPrObPrOThrSerAsnSerArGlnPrObPrObPrOThrAsnPrO 385
QY 9 ---GGCGCGCG 1
    |||
Db 386 TyrGlnArGPrO 389
```

```
RESULT 3
US-09-949-016-7664
/ Sequence 7664, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 7664
/ LENGTH: 324
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-7664

Alignment Scores:
Pred. No.: 0.000883 Length: 324
Score: 126.50 Matches: 52
Percent Similarity: 38.8% Conservative: 12
Best Local Similarity: 31.5% Mismatches: 66
Query Match: 13.7% Indels: 35
DB: 2 Gaps: 9

US-10-071-510A-16 (1-493) x US-09-949-016-7664 (1-324)
QY 440 CCCAGACAGGAGGGGCTGTGCTGAGACCCCTTACCACTTGAACCTGGGCGCTGCG 381
    |||:|||||
Db 125 PrObPrOThrGlnPrObAlaPrObSerGlnPrObPrGlnPrObPrGlnLeuPrObPrObPrO 144
QY 380 CCGTTTCCACAGCCCTGCTAGACAGACTCAGCGCTCGACAGATCCCTGGGAATGATGATC 321
    |||:|||||
Db 145 PrGlnValAlaPrObPrObAlaPrObValAlaPrObPrO-----PrObAlaSerGlnVal 161
```

```
QY 320 CAT-----GGATTGATGAGAAATCTGACTTTGGAGTGTGTGGTGAAC 273
    |||
    |||
    |||
Db 162 H18ProProAlaProGlyValHisProProAla-----ProGlyValHis 176
QY 272 CATTTTCTCCCGAATTCAC-----TTCCATCTCTCGAGATCCAGATT 228
    |||
    |||
    |||
Db 177 ProProAlaProGlyValHisProProThiSerGlyValHisProProAlaProGlyVal 196
QY 227 TCTGATGAGGCTCAGACGACTTTCTTCAGACGCTTCAGAGCCCGCCAGCACTCTGCCAG 168
    |||
    |||
    |||
Db 197 H18ProProAlaProGlyValHis--ProProAla-----ProGlyValHisProProAla 214
QY 167 CTCCTGAGCT---CTCTCTGACACGAGCAGACCTT-----C 132
    |||
    |||
    |||
Db 214 laProGlyValHisProProAlaProGlyValHisProProSerAlaGlyValHisP 234
QY 131 CGAGAAAGATTTCTTCATCCATCCAGCCAGCCGCTTCACAGGAGCAGCTGGG--- 76
    |||
    |||
    |||
Db 234 roGlnAlaProGlyValHisProAlaAlaProAlaValHisProGlnAlaProGlyValHis 254
QY 75 ---CTCTCTCTCCGGAATTTCTGCCACAGCCCTTCMAACTCGGCTCTTGGAGCTCGGC 18
    |||
    |||
    |||
Db 254 laProProAlaProGlyValHisProGlnAlaProGlyValHisPro-----G 270
QY 17 ATCCCCCGGCG 7
    |||
    |||
    |||
Db 270 laProProGly 273

RESULT 4
US-09-252-991A-21715
; Sequence 21715, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21715
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-21715

Alignment Scores:
Pred. No.: 0.001 Length: 339
Score: 126.00 Matches: 64
Percent Similarity: 27.6% Conservative: 13
Best Local Similarity: 22.9% Mismatches: 58
Query Match: 13.7% Indels: 144
Gaps: 2

US-10-071-510A-16 (1-493) x US-09-252-991A-21715 (1-339)
QY 2 GGCCGGCCGGGGG-----TSCCGAGTCCCAAGAGC 34
    |||
    |||
    |||
Db 69 G1YUyrProGlyValHisPProProValHisGlyLeuCybArGAlaAerGlyUGly 88
QY 35 CGAGTTTGAGAGCTGGTGGCAGAAATTCCTCCGAGAGAGGCT----- 76
    |||
    |||
    |||
Db 89 G1YAlaProGlyValHisGlyValAlaProGlyValHisPProAlaGlyGlyAspGln 108
QY 77 -----CCAGCTGTCTCCGTGAGAGCCAGG----- 103
    |||
    |||
    |||
Db 109 ValAlaHisValAlaAlaLeuGlyGlySerArgGlyGlyArgGlyArgSerThrGly 128
QY 103 ----- 103
```

```
Db 129 ArgAlaAerProProProAlaGlyValArgThrValGlnGlyGlyArgArgProArg 148
QY 104 -----CTGCTGTGATGAGAAATCTTCTCCGAGGCTGC 139
    |||
    |||
    |||
Db 149 ArgThrAlaProGlyValArgProArgArgAlaGlyValGly-----GlyGlyCys 164
QY 140 TCCCTGTGTCAGAGAGAGCTCAGGAGCTGCAGAGTCTGTGCGGCGCTTGAGGCTGCT 199
    |||
    |||
    |||
Db 165 SerArgGlyValGlyValGlnAlaAlaGlyValGlyValAlaAlaSerAlaGly 184
QY 200 GGAAGAAAGTCTGCTGAGCT----- 220
    |||
    |||
    |||
Db 185 AerGlyValArgAlaAlaProArgArgProValArgGluProProGlyAspProArgGly 204
QY 221 -----CATCAGAAATCTGCATCTCCAGAGAT----- 247
    |||
    |||
    |||
Db 205 ThrValHisProArgThrHisGlyValSerAlaGlyAlaProGlyValArgArgLeuArg 224
QY 247 ----- 247
Db 225 AlaLeuHisProValValHisProGlyGlyGlnAlaArgArgGlyGlyHisLeuHis 244
QY 248 -----GGAAGTGAATTCGGGAGAGAAATGTTTTCAC 280
    |||
    |||
    |||
Db 245 GlyAspProArgThrGlyGlnGlyValAerGlyGlyThrGlyValGlnArgGlyPhe 263
QY 281 CAACAACATCCCAAGATCCAGATTCTCATCAATCCATGATCTATTCCAGGATCG 340
    |||
    |||
    |||
Db 264 ArgArgHisPro-----GlyArgArgAsnProGlyGlySerProGly 272
QY 341 TCAGCCGCTGAGCTCTTACAGAGGCTGCGAGAGG-----GGCCAGGCCAGGT 394
    |||
    |||
    |||
Db 273 SerProArgSerGly-----GlyArgArgAsnProGlyGlySerProGly 287
QY 395 CA-----GAGGTGGTGGGCTTCAGACAGGCCC----- 427
    |||
    |||
    |||
Db 288 ArgArgPheArgValSerAlaProLeuAlaValArgProProAlaAerGlyArgLeuPhe 307
QY 428 -----CTCCCTGTCTGGGCGCAATGCTCTGTAGAGACTTGCC 469
    |||
    |||
    |||
Db 308 ArgAlaSerGlyPheAlaProCysArgGlyLeuLeuSerCysMetProAerSerAla 326

RESULT 5
US-09-252-991A-22317
; Sequence 22317, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22317
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-22317

Alignment Scores:
Pred. No.: 0.00131 Length: 561
Score: 125.50 Matches: 54
Percent Similarity: 37.3% Conservative: 18
Best Local Similarity: 28.0% Mismatches: 55
Query Match: 13.6% Indels: 66
Gaps: 2
```



```

Qy      170 CAGCTCCC-----TGAGTCCTCTGCACCCAGGCGAGCAGCTCCGGAGAAGA 123
        |||
Db       533 rctrrPrroIngInglyLeuValleuglyPrroProAlaPrroPrroPrroPrroLeuPrroS 553
Qy      122 CTTCATCATCACCAAGCCAGCCCTGCGCTTCCACAGGAGACAGTGGGCTCTCTCCG- 64
        ::|||:::
Db       553 ergIyPrroAltyrAlaser---AlaleuPrroPrroPrroglYPrroPrroPrroP 572
Qy      63 -----GGAATTCTGCCACAGCCTCTCAAATCTGGGCTTTGGAACTCGGC 18
        |||
Db       572 roleuPrroSerThrglyPrroPrroPrroPrroPrroPrroPrroPrroLeuPrroAnGlnA 592
Qy      17 ATCCCCGGGGGGCGCG 1
        |||
Db       592 laPrroPrroPrroPrro 597

RESULT 8
5202236-37
; Patent No. 5202236
; APPLICANT: MAUGH, KATHY J.; ANDERSON, DAVID M.; STRAUSBERG,
; SUSAN L.; MCCADILLIS, RUSS; WEI, TENA; FILIPULA, DAVID
; TITLE OF INVENTION: METHOD OF PRODUCING BROADHESIVE
; PROTEIN
; NUMBER OF SEQUENCES: 39
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/528,762
; FILING DATE: 25-MAY-1990
; APPLICATION NUMBER: 82,456
; FILING DATE: 07-AUG-1987
; APPLICATION NUMBER: 933,945
; FILING DATE: 24-NOV-1986
; APPLICATION NUMBER: 650,128
; FILING DATE: 13-SEP-1984
; SEQ ID NO:37:
; LENGTH: 331
5202236-37

Alignment Scores:
Pred. No.:          0.0024          Length:          331
Score:             122.00           Matches:         26
Percent Similarity: 41.6%            Conservative:    6
Best Local Similarity: 33.8%          Mismatches:     33
Query Match:       13.2%              Indels:        12
DB:                6                  Gaps:          1

US-10-071-510A-16 (1-493) x 5202236-37 (1-331)
Qy      261 CCGAATCCACTTCATCTCTGCAGATGCCAGTTCTGTATGAGGCTCAGCAGACTTTCTT 202
        |||
Db       21 PrroAnPrroAnPrroPrroSerPro----- 28
Qy      201 CCAGCAGCCTCAAGGCCCGCCGACGACTGTCCAGCTCCTGAGCTCTCTCGACACGAG 142
        |||
Db       29 ProSerProPrroSerProPrroThPrroPrroThPrroPrroSerProPrroAlaPrroP 48
Qy      141 CAGCACCCTCCGAGAGAAGACTTCTCCATACACGACGAGCCTGCGCTTCACACGAGAGA 82
        :|||:
Db       49 SerPrroPrroPrroSerProPrroAnPrroPrroPrroSerProPrroPrroSerProPrroPhPe 68
Qy      81 GCTGGGCTCTTCTTCGGGAAATTCGCCACACGAGCCTCTCAAACTGGAGCT 31
        |||
Db       69 PrroAlaPrroPrroSerProPrroPhPePrroPrroThTylysAlaLySp 85

RESULT 9
US-09-949-016-9978
; Sequence 9978, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0001307

```

```

CURRENT APPLICATION NUMBER: US/09/949, 016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241, 755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237, 768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231, 498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9978
LENGTH: 581
TYPE: PRT
ORGANISM: Human
US-09-949-016-9978

Alignment Scores:
Pred. No.: 0.00399 Length: 581
Score: 120.50 Matches: 47
Percent Similarity: 33.9% Conservative: 9
Best Local Similarity: 28.5% Mismatches: 50
Query Match: 13.1% Indels: 59
DB: 2 Gaps: 7

US-10-071-510A-16 (1-493) x US-09-949-016-9978 (1-581)
QY 467 CCAAGTCTCTGAGAGCAGCATGTTGCCCCAGACAGGAGGGGCGCTGTGCTGAGACCCC 408
DB 355 ProProProProSerArg-----GlyGlyProProProProPro 367
QY 407 TACCACCTCTTGACCTGGGGGCTGGCCCTTTCTCCACAGCCCTGTAGACAGATCC 348
DB 368 ProProProProHisSerSerGlyProProPro-----ProProAlaArgGlyArgGly 384
QY 347 GCGTCAGACGATGCGCTGGGATAGATCCATGGGATGTGAGAAATCCTGACTTTGGAT 288
DB 385 AlaProProProProProSerArg-Ala----- 393
QY 287 GTTGTGTGAAAAACCATTTTCTTCCCGAATCCATTCATCCTCTGCAGATGCCAGTT 228
DB 394 -----ProThrAlaAlaProProPro----- 400
QY 227 TCTGATGAGGCTCAGACAGACTTTCTTCAGACGCTCAAGGCCGCCACGACTTGCACG 168
DB 401 -----ProProProSerArgProSerValAlaValPro- 411
QY 167 CTCCCTGAGTTCCTCTGTCACACGCGACGACCTTCGGAGAAAGATTCTCATCACAG 108
DB 412 -----ProProProProAlaGlyMetGlyProPro-----ProProPro 425
QY 107 CCAGCCCTGCGGCTTCACACGAGGAGCAGCTGGGCTCTCTTCGGGGAATTCG----- 55
DB 425 lAlaProSerSerAlaProSerGlyProProProProProProSerValAlaGlyValG 445
QY 54 -----CCACGAGCTCTCAAACTCGGCTCTGGGAGCATCGGATCCCGC 12
DB 445 lProValAlaProProProProProProProProProProGlyProProProPro 465
QY 11 CGGGCGGGCG 1
DB 465 foGlyLeuPro 468

RESULT 10
US-09-477-962-107
Sequence 107, Application US/09477962
Patent No. 6927286
GENERAL INFORMATION:
APPLICANT: SHEN, BEN
APPLICANT: DU, LIANGCHENG
APPLICANT: SANCHEZ, CESAR
APPLICANT: CHEN, MEI
APPLICANT: EDWARDS, DANIEL J.
TITLE OF INVENTION: BLEBOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES

```

```

; FILE REFERENCE: 4071-895820US
; CURRENT APPLICATION NUMBER: US/09/477,962
; CURRENT FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: 60/115,435
; PRIOR FILING DATE: 1999-01-06
; PRIOR APPLICATION NUMBER: 60/118,848
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 107
; LENGTH: 935
; TYPE: PRT
; ORGANISM: Streptomyces verticillius
; FEATURE:
; OTHER INFORMATION: ORF16
US-09-477-962-107

Alignment Scores:
Pred. No.: 0.00464 Length: 935
Score: 120.50 Matches: 56
Percent Similarity: 37.2% Conservative: 18
Best Local Similarity: 28.1% Mismatches: 59
Query Match: 13.1% Indels: 66
                Gaps: 11
US-10-071-510A-16 (1-493) x US-09-477-962-107 (1-935)

QY 436 GACAGGAGGAGGCGCTGTGCTGAGAGACCCCTTACCCACTTGTGACTGGGCGCTGAGCCCT 3777
Db 508 Asphargylarhyvalhrhlaaglualaprohlaalavalprohlaalapro 5272
QY 376 TCTCCCAAGACC-----CTGCTAGACAGACTACCGCTGCACGA 338
Db 528 AAlaserhlaProhlaargapglualgluleuvalaglnvalaIargvalthr 5478
QY 337 TGC-----CTGGGAATAGGATCCATGGGAGATTGATGAGAAATCTGACTTTGGATGTTG 2844
Db 548 Cyahrghvalenuglylleglyalaval-----GluProaspMetasnuleu 5635
QY 284 -----284
Db 564 AsplaclaylahrthrservalGluLeuValArgLeuAlathrAlaLeuGluGluLeu 5835
QY 283 -----TTGGTGAAMACCATTTTCTTCCCGAATCCATCCATCCATCTCTGC 239
Db 584 GlyLeuaspThrasplelleglugluLeuAlaPhePro--SerValaIaValIleVal 6020
QY 238 AGATGCAGCTTTGTGATGGAGCTCACA-----GACTTCTTCCAGCAGCTC 191
Db 603 GLyArghlsleucllyargrg-ThrhlaprohlaIarghapProleuProProhlaIse 6222
QY 190 A-----AGCCCGCCACAGACTTGCACACTCCCTGAGTCTCTCTGCACACGGCA 140
Db 622 rValaIaPheAlaProglYserValleuPro-----AlaProProhlaProglYPr 639
QY 139 GCACCCCTCCGAGAGAAGACTTCTTCATCACCACGACAG-----CCTGGGCTTCCACAGG 86
Db 639 oValProhlaIaservalProhlaIaProhlaIaservalProhlaIasergIuserse 659
QY 85 GACAGCTGGGCGCTCTTCTCCGGA-----ATTCT 56
Db 659 rProleuAlaProhlaIaProglYProvalProhrothProvalProhlaIaserva 679
QY 55 GCCACGAGCGCTTCaAACTCGGCTTTGGAGCTGGGATCCCGGGCGGCGG 1
Db 679 lProProhlaIaserglyAlaIaIaPro-----HlvalProhlaIaPro 693

RESULT 11
; US-08-618-112-143
; Sequence 143, Application US/08018112
; Patent No. 6290969
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.

```

```

? APPLICANT: Skeiky, Yasir A.W.
? APPLICANT: Dillon, David C.
? APPLICANT: Campos-Neto, Antonio
? APPLICANT: Houghton, Raymond
? APPLICANT: Vedvick, Thomas S.
? APPLICANT: Wardzik, Daniel R.
? TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
? NUMBER OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS
? NUMBER OF SEQUENCES: 153
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: SEED AND BERRY LLP
? STREET: 6300 Columbia Center, 701 Fifth Avenue
? CITY: Seattle
? STATE: Washington
? COUNTRY: USA
? ZIP: 98104-7092
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/818,112
? FILING DATE: 13-MAR-1997
? CLASSIFICATION: 424
? ATTORNEY/AGENT INFORMATION:
? NAME: Maki, David J.
? REGISTRATION NUMBER: 31,392
? REFERENCE/DOCKET NUMBER: 210121.411C6
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (206) 622-4900
? TELEFAX: (206) 682-6031
? INFORMATION FOR SEQ ID NO: 143:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 174 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-08-818-112.143

Alignment Scores:
Pred. No.: 0.00305 Length: 174
Score: 120.00 Matches: 45
Percent Similarity: 31.1% Conservative: 2
Best Local Similarity: 29.8% Mismatches: 56
Query Match: 13.0% Indels: 48
DB: 2 Gaps: 8

US-10-071-510A-16 (1-493) x US-08-818-112-143 (1-174)

QY 411 CCCCTACCCACCTCTTGAGACTTGGGGCCCTGG-----CCCTTCTCCACAGCCCTGC 361
|||
Db 29 ProSerProProSerProProthrgLyTyrPvalProArgalalauLeuPro----- 45
360 TAGACAGACTACAGCGGTGCAGCAGATGCTGGGAATAGATCCATGAGATTGATGAGAAATC 301
46 -----ProTPr----- 47
QY 300 CTGACTTGGAGATGTTGGTGTAAMACCATTTCCTCCCGGAATCCACTTCCATCTCT 241
|||
Db 48 ---LeuallagLyThrProProalalProProvalProPrometalalProleuPro----- 64
240 GCAGATGCCAGATTCTGTAGTAGAGCTCAGCAGACATTTCTTCAGCAGGCTCAAGGCCGCC 181
65 -----ProalalalProleuPro--- 70
180 ACGACTTGCACGCTCCCTGAGCTTCCTCTGCACACGAGCAGCACCTCCGAGAGAACT 121
|||
QY 71 -----ProleuProProleuProProleuProthrsrHisProProalrProPro 87
:::|||||
120 TCTCCA-----TACCGACGCCAGCCCTTGG-----CTTCACCCAGGGACA 82
:::|||||

```

Db 88 AlapProAlaProProAlaProAlaCysProPheValProValProProAlaPro 107  
QY 81 GCTGGGCTCTCTCTCG-----GGAATTCTGCCACGCTCT 43  
Db 108 ProleuProProSerProProThrgluLeuProAlaAspAlaAlaCysProProAlaPro 127  
QY 42 CAAACTCGGCTTTGGGACTCGGCATCCCGC 10  
Db 128 ProAlaProProLeuAlaProProSerProPro 138

RESULT 12  
US-08-818-111-138  
Sequence 138, Application US/08818111  
Patent No. 6338852  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
APPLICANT: Campos-Neto, Antonia  
APPLICANT: Houghton, Raymond  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Twardzik, Daniel R.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 148  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,111  
FILING DATE: 13-MAR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.4176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 138:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-818-111-138

Alignment Scores:  
Pred. No.: 0.00305 Length: 174  
Score: 120.00 Matches: 45  
Percent Similarity: 31.1% Conservative: 2  
Best Local Similarity: 29.8% Mismatches: 46  
Query Match: 13.0% Indels: 58  
DB: 2 Gaps: 8

US-10-071-510a-16 (1-493) x US-08-818-111-138 (1-174)  
QY 411 CCCCTACCCACTTTGACTGGGGCTCG-----CCCCTTCTCCACAGCCCTGC 361  
Db 29 ProSerProProSerProProThrglyTyrValProAlaAlaLeuLeuPro----- 45  
QY 360 TAGACAGACTCAGCGCTGCAGATGCTGGGAATAGATCGAGATTGATGAGAAATC 301  
Db 46 -----ProTrr----- 47

QY 300 CTGACTTTGGAGTGTGTGTGGTGAACCATTTCTCTCCCGAATCCATTCCTCT 241  
Db 48 ---LeuAlaGlyThrProProAlaProProValProProMetAlaProleuPro----- 64  
QY 240 GCAGATGCCAGTTTCTGATGAGCTCAGCAGACTTTCTTCAGCAGCCTCAAGCCCGCC 181  
Db 65 -----ProAlaAlaProleuPro----- 70

QY 180 ACGACTTGGCCAGCTCTCTGAGCTCCTCTGCAACGACGACGACCCCTCGGAGAAAGAT 121  
Db 71 -----ProleuProProleuProProleuProThSerHisProProArgProPro 87

QY 120 TCTCCA-----TCACGAGCAGCCCTCG-----CTTCGACGAGGACA 82  
Db 88 AlapProProAlaProProAlaProProAlaCysProPheValProValProProAlaPro 107

QY 81 GCTGGGCTCTCTCTCG-----GGAATTCTGCCACGCTCT 43  
Db 108 ProleuProProSerProProThrgluLeuProAlaAspAlaAlaCysProProAlaPro 127  
QY 42 CAAACTCGGCTTTGGGACTCGGCATCCCGC 10  
Db 128 ProAlaProProLeuAlaProProSerProPro 138

RESULT 13  
US-09-056-556-143  
Sequence 143, Application US/09056556  
Patent No. 6350456  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND  
NUMBER OF SEQUENCES: 241  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/056,556  
FILING DATE: 07-APR-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Makl, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.457  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 143:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-056-556-143

Alignment Scores:  
Pred. No.: 0.00305 Length: 174  
Score: 120.00 Matches: 45  
Percent Similarity: 31.1% Conservative: 2  
Best Local Similarity: 29.8% Mismatches: 46  
Query Match: 13.0% Indels: 58

DB:	2	Gaps:	8
US-10-071-510A-16 (1-493) x US-09-056-556-143 (1-174)			
Qy	411 CCCCTACCCACCTCTTGAGCTGGGGGCTGG-----CCCTCTTCCACAGCCCTGC	361	
Db	29 ProberProProberProProThnGlyTrrValProArgAlaLeuLeuPro-----	45	
Qy	360 TAGACAGACTCAGCCGCTCGACAGTGCCTGGGAATAGATTCATCGAGATTGATGAGAAATC	301	
Db	46 -----ProTrr-----	47	
Qy	300 CTGACTTTGGGATGTTGTGGTGAACAATTTTCTTCCCGAATCCACTTTCATCTCT	241	
Db	48 ---LeuAlaGlyThrProProAlaProProValProPrometAlaProLeuPro-----	64	
Qy	240 GCAGATGCCAGTTTCTGATGAGGCTCAGACAGACTTTCTTCACAGAGCTTCAGGCCGCC	181	
Db	65 -----ProAlaAlaProLeuPro---	70	
Qy	180 ACGACTTGCCAGCTCCCTGAGCTCTCTTCGACACGACGACGACACCTTCGGAGAAACT	121	
Db	71 -----ProLeuProProLeuProProLeuProThrsrHisProProArgProPro	87	
Qy	120 TCTGCA-----TCAGCAGCAGCCCGCG-----CTTCACACGAGGACA	82	
Db	88 AlaProProAlaProProAlaProProAlaCysProPheValProValProProAlaPro	107	
Qy	81 GCTGGGCTCTCTCTCCG-----GGAATTCTGCCACGAGCTCT	43	
Db	108 ProLeuProProSerProProThrGluLeuProAlaPheAlaAlaCysProProAlaPro	127	
Qy	42 CAAACTGGCCTCTTGGGACTCGGCATCCCCG	10	
Db	128 ProAlaProProLeuAlaProProSerProPro	138	
RESULT 14			
US-09-072-596-138			
Sequence 138, Application US/09072596			
Patent No. 6458366			
GENERAL INFORMATION:			
APPLICANT: Reed, Steven G.			
APPLICANT: Skeiky, Yaser A.W.			
APPLICANT: Dillon, Davin C.			
APPLICANT: Campos-Neto, Antonia			
APPLICANT: Houghton, Raymond			
APPLICANT: Vedvick, Thomas S.			
APPLICANT: Twardzik, Daniel R.			
APPLICANT: Lodes, Michael J.			
APPLICANT: Hendrickson, Ronald C.			
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF			TUBERCULOSIS
NUMBER OF SEQUENCES: 350			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: SEED and BERRY LLP			
STREET: 6300 Columbia Center, 701 Fifth Avenue			
CITY: Seattle			
STATE: Washington			
COUNTRY: USA			
ZIP: 98104-7092			
COMPUTER READABLE FORM:			
MEDIUM TYPE: floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/072.596			
FILING DATE: 05-MAY-1998			
CLASSIFICATION:			
ATTORNEY/AGENT INFORMATION:			
NAME: Makl, David J.			
REGISTRATION NUMBER: 31,392			
REFERENCE/DOCKET NUMBER: 210121.417C9			
TELECOMMUNICATION INFORMATION:			

```

:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ. ID NO: 138:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 174 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
: US-09-072-596-138
:
:
:
: Alignment Scores:
: Pred. No.:
:
: Score: 0.00305 Length: 174
: Percent Similarity: 120.00 Matches: 45
: Best Local Similarity: 31.1% Conservative: 2
: Query Match: 29.8% Mismatches: 46
: DB: 13.0% Indels: 58
: Gaps: 8
:
:
: US-10-071-510A-16 (1-493) x US-09-072-596-138 (1-174)
:
: Oy 411 CCCCTACCCACCTCTTGTACCTGGGGCCCTG-----CCCTCTTCCACAGCCCTGC 361
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 29 ProSerProProSerProSerProThrGlyIryPalProArgalaleuPro----- 45
:
: Oy 360 TAGACAGACTCAGCGGTGCAGCATGCCCTGGGAATAGATTCATGGATGATGGAATC 301
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 46 -----ProTrp----- 47
:
: Oy 300 CTGACTTTGGAGATGTTGTTGGTGAACCATTTTCTTCCCGAATCCACTTCCTCT 241
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 48 ---LeuAlaGlyThrProProAlaProProAlaProProMetAlaProleuPro----- 64
:
: Oy 240 GCAGATGCAGTTTCTGATGAGCTCAGCAGACTTTCTTCCAGCAGCCTCAAGGCCGCC 181
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 65 -----ProAlaAlaProleuPro--- 70
:
: Oy 180 ACGACTCTGCCAGCTCCCTGAGCTCCTCTGCACACAGCGACCCCTCGGAGAAAGCT 121
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 71 -----ProleuProProleuProProleuProThrSerHisProProAlaArgProPro 87
:
: Oy 120 TCTCCA-----TCACGACCCAGCCCTCG-----CTTCCACCGAGGACA 82
: :||: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 88 AlaProProAlaProProAlaProProAlaIaCysProPheValProValProProAlaPro 107
:
: Oy 81 GCTGGGCTCTCTTCCG-----GGAATTCTGCCACGACCTCT 43
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 108 ProleuProProSerProProThrGluLeuProAlaIaPheAlaIaCysProProAlaPro 127
:
: Oy 42 CAAACTGGCCTTTGGGACTGGGATCCCGCC 10
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 128 ProAlaProProleuAlaProProSerProPro 138
:
:
: RESULT 15
: US-09-072-967-143
: ; Sequence 143, Application US/09072967
: ; Patent No. 6592877
: ; GENERAL INFORMATION:
: ; APPLICANT: Reed, Steven G.
: ; APPLICANT: Skeiky, Yasir A.W.
: ; APPLICANT: Dillon, Davin C.
: ; APPLICANT: Campos-Neto, Antonio
: ; APPLICANT: Houghton, Raymond
: ; APPLICANT: Vedvick, Thomas S.
: ; APPLICANT: Twardzik, Daniel R.
: ; APPLICANT: Lodes, Michael J.
: ; APPLICANT: Hendrickson, Ronald C.
: ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
: ; NUMBER OF SEQUENCES: 355
: ; CORRESPONDENCE ADDRESS:
: ; ADDRESSEE: SEED and BERRY LLP
: ; STREET: 6300 Columbia Center, 701 Fifth Avenue
:

```





**THIS PAGE LEFT BLANK**



Score:	156.50	Matches:	47
Percent Similarity:	38.4%	Conservative:	6
Best Local Similarity:	34.1%	Mismatches:	34
Query Match:	17.0%	Indels:	51
DB:	5	Gaps:	6

US-10-071-510A-16 (1-493) X US-10-450-763-52278 (1-325)

Qy	408	CTACCCACCTCTTGAACCTGGGGGGCTGGGACCCCTTCTCCGACAGCCGCTGTGACAGACTCA	349
		:::	
Db	101	LeuProProProProProProAlaSerProProLeuLeuPro-----	113
Qy	348	CGCGTCGACGATCCTGGGAATAGATCCATGGGATGTATGAGAAATCTGACTTTGGGA	289
Db	113	-----	113
Qy	288	TGTTGTGGTAAAAACCATTTTCTTCTCCCGAATCCACTTCGATCCTCTGACGATGCCAGT	229
Db	114	-----ProAlaLeuSerPro-----ProLeuProAlaProPro-----	124
Qy	228	TTCTGATGAGGCTCGACGACACTTTCTTCCAGAGCCGTCGAAGGGCCGGCAGACCTGCGCA	169
Db	125	-----ProProSerAlaProProAlaSerProProPro	135
Qy	168	GCTCCCTGAGCTCTCTCTGTCACACGACGGACGACCTCTCCGGAGAGAAGCTTCTTCATCA--	112
Db	136	AlaProProProProProProAlaProProProSerProProProAlaProProProSerAla	155
Qy	111	CCAGCGACGCTGCG--CTTCCACGAGGAGACGCTGGGCTCTCTTCTCCGGAAATTCTG	55
Db	156	ProSerSerProAlaProLeuProProAlaProAlaSerProProSerProAlaProPro	175
Qy	54	CCACGAGCCTCTCAAACTGGGCTCTTGGGAGCTGGGACATCCCGGGCGGCGCG	1
Db	176	ProProAlaProProProAlaPro-----HisProProSerProPro	189

RESULT 2

```

US-10-437-963.139610
: Sequence 139610, Application US/10437963
: Publication No. US20040123343A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: APPLICANT: Wu, Wei
: APPLICANT: Boukharov, Andrey A.
: APPLICANT: Bardazuk, Brad
: APPLICANT: Li, Ping
: TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(53221)B
: CURRENT APPLICATION NUMBER: US/10/437,963
: CURRENT FILING DATE: 2003-05-14
: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 139610
: LENGTH: 220
: TYPE: PRT
: ORGANISM: Oryza sativa
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(220)
: OTHER INFORMATION: unsure at all Xaa locations
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT4530_40885C.1.pep
: US-10-437-963-139610

```

Alignment Scores:	
Pred. No.:	0.00119
Score:	149.50
Percent Similarity:	33.5%
Best local Similarity:	28.7%
Query Match:	16.2%
Length:	2220
Matches:	47
Conservative:	8
Mismatches:	58
Indels:	51

DB:	4	Gaps:	6
US-10-071-510A-16 (1-493) x US-10-437-963-139610 (1-220)			

Oy	411	TTGAGAGCCCTACCCACCTCTTGACCTGGGGGCTGGCCCTTTC-	372
Db	76	TTPCyEProProProLensePProProCySProTTPProTTP**CyEProProPro	95
Oy	372	-----CCACAGCCCTGTAGACAGACTCAGCGCTGCAGCATGCTGGGAAATAGCAT	322
Db	96	ProProProProGlnProProProProProProGlnProProProProTTPSerProGlu	115
Oy	321	CCATGGAGATTGATGAGAAATCCTGACTTTGGGATGTGT-----TGGTGAAA	274
Db	116	LYeTTP-----CySProProProProTTPSerPro	125
Oy	273	CCATT-----TCTTC	262
Db	126	ProTTPTrpCyEglnProProPro**SerProProCySerTTPCySProProSer	145
Oy	261	CCGAAATCCACTTCACCTCTGACAGATGCGAGTTTGATAGGCTCAGCATTTCTT	202
Db	146	ProProProSerProProSerSerSerSerSerSerSerSerSerSerSerSerSerSer	156
Oy	201	CCAGCAGCCTCAAGGCCCGCCACGACTGTGCAGCTCCCTGAGCTCTCTGCACGCG	142
Db	157	ProProProSerProPro-----ProProProSerAlaProProProProSer	172
Oy	141	CAGACCTCCGGAGAAAGCTTCTCCATCACAGCAGCACCCTGGGGCTTCCACCGAGGACA	82
Db	173	AlaProProProGlnProSerAlaProProProSerProProSerSerProProPro	122
Oy	81	GCTGGGCTCTCTTCTCCGGAATTCTGCCACCGAGCTCTCAAACTCGGCTTTGGAGCT	22
Db	193	ProProProProSerSerProProProSerProProAlaProGlnProSerLeuAnleu	212
Oy	21	CGGATCCCGG	10
Db	213	AlaValAlaPro	216

### RESULT 3

```

US-10-437-963-116465
; Sequence 116465, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221) B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 116465
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(263)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4550_19964C.1 pep
; US-10-437-963-116465

```

Alignment Scores: 0.00163 Length: 263  
Pred. No.:

Score:	144.50	Matches:	40
Percent Similarity:	38.4%	Conservative:	12
Best Local Similarity:	30.5%	Mismatches:	56
Query Match:	15.7%	Indels:	38

Score:	44.00	Conservative:	10
Percent Similarity:	32.7%	Mismatches:	43
Best Local Similarity:	27.6%	Indels:	89
Query Match:	15.7%		

DB:	4	Gaps:	8
DB:	US-10-071-510A-16 (1-493)	x US-10-437-963-168762 (1-437)	
Qy	467 CCAAGTCTCAGAGCAGCATGT--TGCCCCAGACGAGGAGGCGCTGTGTGAGAC	411	
Db	225 PromerProProProSerSerAlaAlaProProSerSerGlyGlyArgCysAla-----	242	
Qy	410 CCCTACCCACCTTTGACCTGGGGCCCTGCCCCCTTTCGCCACAGCCCTGTGCACAGACT	351	
Db	243 -----Ser--GlySerSerProArgArgSerProProCys-----	253	
Qy	350 CACGCGTCGACGATGCTGGGAATAGGATCATGATGATGATGAGAAATCCTGACTTTGG	291	
Db	253 -----	253	
Qy	290 GATGTGTGTGTGMAAACATTTTCTTCCCGAATCAGCTTCATCCTGTGAGATGCA	231	
Db	254 -----SerProSerThSerCysSerProSerProSerSer-----	265	
Qy	230 GTTTCGTGATGAGGCTCAGCAGACTTTCTTCAGACAGCCTCAAGGCCGCCACGACT--C	174	
Db	266 -----ProAlaAlaAlaThrThSerThrThrArgC	276	
Qy	173 TGCCAGCTCCCTGAGGCTCCTCGACACGAG-----CAGC	138	
Db	276 yseProleuProGluGlnProProSerProhiStrh**LeuSerProArgAlaAlaSerP	296	
Qy	137 ACCCTCCGGAAGAACTTCTCCA-----	115	
Db	296 roProProSerProSerSerProProProProCysGlySerProleuProGlySerSers	316	
Qy	114 -----TCACAGCCAGCCCTCGGCTTCACACGAGGACG	81	
Db	316 eSerProArgThrProProSerSerSerProAlaSerProAlaIleProProSerProA	336	
Qy	80 CTGGGCGCTCT-----TCCTCGGGAATTCGTGCACAGCCT	45	
Db	336 laProProProhiSleuLeuSerSerProProGlySerProSerProvalProleuP	356	
Qy	44 CTCAAACTCGGCGCTTGGGACTCGGCATCCCCCGGCGCCGCG	1	
Db	356 roProProProProleuGlySerThrArgArgProProSerPro	370	
RESULT 6	US-10-437-963-189874		
	Sequence 189874, Application US/10437963		
	Publication No. US20040123343A1		
	GENERAL INFORMATION:		
	APPLICANT: La Rosa, Thomas J.		
	APPLICANT: Kovalic, David K.		
	APPLICANT: Zhou, Yihua		
	APPLICANT: Cao, Yongwei		
	APPLICANT: Wu, Wei		
	APPLICANT: Boukharov, Andrey A.		
	APPLICANT: Barbazuk, Brad		
	APPLICANT: Li, Ping		
	TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With		
	TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement		
	FILE REFERENCE: 38-21(53221)B		
	CURRENT APPLICATION NUMBER: US/10/437, 963		
	CURRENT FILING DATE: 2003-05-14		
	NUMBER OF SEQ ID NOS: 204966		
	SEQ ID NO 189874		
	LENGTH: 148		
	TYPE: PRT		
	ORGANISM: Oryza sativa		
	FEATURE:		
	NAME/KEY: unsure		
	LOCATION: (1)..(148)		
	OTHER INFORMATION: unsure at all Xaa locations		
	FEATURE:		
	OTHER INFORMATION: Clone ID: PAT_MRT4530_86340C.1 pep		

```

US-10-437-963-189874

Alignment Scores:
Pred. No.: 0.0038      Length: 148
Score: 143.00         Matches: 36
Percent Similarity: 41.2%      Conservative: 6
Best Local Similarity: 35.3%   Mismatches: 38
Query Match: 15.5%           Indels: 22
DB: 4                      Gaps: 4

US-10-071-510A-16 (1-493) x US-10-437-963-189874 (1-148)
OY      282 TGGTGAACAATTTCTTCCCGAATCCACTTCCTGCAGATGCCAGTTTCTGA 223
      ||| :||| ||||| :||| ||| :|||
DB      45 TTPAAGAGProSerSerSerSerSerProThProAspAla**Pro----- 60
OY      222 TGAGGCTCAGCAGACTTCTTCCAGCAGGCTCTCAAGGCCCGGCCACGACTTCGACGTTCCC 163
      ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      61 -----ProAspAla**ProProSerProAlaProProPro 73
OY      162 TGAGTCTCTCTGCACACGCGACGACCTTCGCGAAGAAGACTTCTCA----- 115
      ||||| ||||| ||||| ||||| ||||| |||
DB      74 ArgThrProSerProSerProSerProProProProProSerProProProProPro 93
OY      114 -----TCACGACGACGCGCTTGCGCTTCACACGAGGACAGCTGGGCGCTTCT 67
      ||| ||| ||||| ||||| ||||| ||||| |||
DB      94 ThrGlyAsp**ProProProSerProValLeuProProArgThr-----ProProSer 111
OY      66 CCGGAATTTCTGCCACGACGCTCTCAAGCTCGGCTCTTGGAAGCTCGGCATCCCGCGAGC 7
      ||| ||||| ||||| ||| ||| ||||| |||||
DB      112 ProSerProAlaProProProProProSerProAlaPro---ProProProSerProProGly 130
OY      6 CGGCGCG 1
      |||||
DB      131 ArgPro 132

RESULT 7
US-10-437-963-161137
; Sequence 161137, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 161137
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(501)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4520_60349C.1.pep
US-10-437-963-161137

Alignment Scores:
Pred. No.: 0.00561      Length: 501
Score: 142.00         Matches: 51
Percent Similarity: 35.7%      Conservative: 9
Best Local Similarity: 30.4%   Mismatches: 54
Query Match: 15.4%           Indels: 54

```

```

DB: 4 Gaps: 6
US-10-071-510A-16 (1-493) x US-10-437-963-161137 (1-501)
QY 476 AGAGCGTGGCCAAAGTCTTCAGAGCAGAGATGTGGCCCAAGAGGAGGGGCTGTGCT 417
   |||||:|||||
Db 108 ATGTTPhneProProProser-----Pro 116
QY 416 GGAGACCCCTACCCACTTGTGACTGTGGGCGCTGCGCCCTTCTCCAGAGCCCTGTAGA 357
   |||||:|||||
Db 117 ProProProAProProProProProProProProProProProProProProProPro 136
QY 356 CAGACTCAGCGCTGACGATGCTGGGAATAGGATCCATGGCATGATGAGAAATCTCTGA 297
   |||||:|||||
Db 137 ThThr**ProProThSerProProPro--ProProThr-----149
QY 296 CTTTGGGATGTTGTGTGTAAGAAACATTTCTCCCGAATCCACTTCCTCTGAG 237
   |||||:|||||
Db 150 -----ProProThGlnProSerAnProLeuProAlaPro----161
QY 236 ATGCCAGTTTGTATGAGGCTCAGACACTTCTTCACGAGCGCTGAAGCGCCGCGCAGA 177
   |||||:|||||
Db 162 -----ProSerSerSerProProProThrl 170
QY 176 CTCTGCCAGCTCCCTGAGCTCTCTCTGACACGAGCAGACCCCTCGGAGAAACTTCTC 117
   |||||:|||||
Db 170 hProProProProProProProProProProProProProProProThSerP 185
QY 116 CATCACGAGCAGCCCTGCGCTCCACAGGAGCAGCTGGGCGCTCTCTC-----67
   |||||:|||||
Db 185 roserThrProAPro**ProProProSerSerSerProProThThProProT 205
QY 66 -----CCGGGAATTCGCGACAGCTCTCTCAACTCGGCTCTTGAGACTC 21
   |||||:|||||
Db 205 hserAnProProProProProProThProProSerSerSerProPro-----Thrw 223
QY 20 GGCATCCCCCGGCGCGCG 1
   |||||:|||||
Db 223 eProProProSerThPro 229

RESULT 8
US-10-425-115-289952
; Sequence 289952, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 289952
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(224)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_27523C.1.Dep
; US-10-425-115-289952

Alignment Scores:
Pred. No.: 0.00593 Length: 224
Score: 141.00 Matches: 43
Percent Similarity: 34.1% Conservative: 53
Best Local Similarity: 31.2% Mismatches: 53
Query Match: 15.3% Indels: 38

```

```

DB: 4 Gaps: 5
US-10-071-510A-16 (1-493) x US-10-425-115-289952 (1-224)
QY 414 AGACCCCTACCCAGCTTGTGACTGGGGCGCTGCGCCCTTCTCCACAGCCCTGTAGACA 355
   |||||:|||||
Db 38 AAgProLeuProLeuProAlaPro--AlaProProPro**ProProProAla-----54
QY 354 GACTCAGCGCTGACGATGCTCGGAATAGGATCCATGGGATTTGATGAGAAATCTGACT 295
   |||||:|||||
Db 55 -----Pro-----55
QY 294 TTGGGATGTTGTGGTGAAGAAACATTTCTCCCGAATCCACTTCATCTCTGAGAT 235
   |||||:|||||
Db 56 -----ProPro**ProProGluProLeuProAlaProAla-----68
QY 234 GCCAGTTTGTATGAGGCTCAGACACTTCTTCCAGACGCTCAAGCGCCGCGCAGACT 175
   |||||:|||||
Db 69 -----ProProProTyProProProAlaProProAla-----80
QY 174 CTGCCAGCTCCCTGAGCTCTCTCTGACACGAGACACCTTCGAGAAACTTCTCA 115
   |||||:|||||
Db 81 ---ProProPro**ProProProAlaProProProTyProProProLeuProProPro 99
QY 114 TCACGAGCAGCGCTGCGCTTCACAGGAGAGAGCTGGGCGCTCTCTCGGGAATTCTG 55
   |||||:|||||
Db 100 AlAProAlaProProPro**ProProProAlaProAlaProProProProProAPro 119
QY 54 CCACGAGCTCTCAAACTCGGCTCTTGGAATCGGCAATCCCGGAGCGCGCG 1
   |||||:|||||
Db 120 ProProProProSerProProProProAlaProProGluProProProAPro 137

RESULT 9
US-10-425-115-263984
; Sequence 263984, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 263984
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(231)
; OTHER INFORMATION: Clone ID: MRT4577_172369C.1.dep
; US-10-425-115-263984

Alignment Scores:
Pred. No.: 0.00596 Length: 231
Score: 141.00 Matches: 39
Percent Similarity: 35.3% Conservative: 10
Best Local Similarity: 28.1% Mismatches: 46
Query Match: 15.3% Indels: 44
DB: 4 Gaps: 4

US-10-071-510A-16 (1-493) x US-10-425-115-263984 (1-231)
QY 411 CCCTACCCGACCTGTGACTGGGGCGCTGCGCCCTTCTCCACAGCCCTGTAGACAGAC 352
   |||||:|||||
Db 62 ProAlaProThrAlaThrProProProAlaProThrProProProAlaProThr----80
QY 351 TCACGCGCTGACGATGCTCGGAATAGGATCATGGATTTGATGAGAAATCTGACTTTG 292
   |||||:|||||
Db 80 -----80

```

```
QY 291 GGATGTTGTGGTGAACCAATTTCTCCCGAATCCATTCCATCTCTCGAGATGCG 232
Db 81 -----ThrProProAlaProThrProProPro 91
QY 231 AGTTCTGATGAGGCTCAGCAGACTTTCTTCAGCAGCCTCAAGGCCGCCAGACTGTG 172
Db 92 Ala-----ProThrThProProProAlaProThrTh 102
QY 171 CCAGCTCCCTGAGCTCTCTCGACAGCGGACGCCCTCCGAGAAAGCTTCATCA 112
Db 103 ProProProSerProProAlaSerPro-----ProProAlaProThrThProPro 119
QY 111 CCAGCAGCGCTGCGCTTCACACGAGGAGAGCTGG-----CCTCTCTTCGCGGAAT 58
Db 120 ProSerProProAlaSerProProProAlaProAlaThProProProSerProPromet 139
QY 57 CTGCCACGAGCCTCTCAAACTCGGCTCTTGGACTCGGCATCCCGGCGCGCG 1
Db 140 AlaProProProAlaThProProProProAlaThProProProProAlaAlaPro 158

RESULT 10
US-10-739-930-10816
; Sequence 10816, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 10816
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(238)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: TRIAE-23ABR03-C730_1224.p
US-10-739-930-10816

Alignment Scores:
Pred. No.: 0.00599 Length: 238
Score: 141.00 Matches: 39
Percent Similarity: 42.6% Conservative: 7
Best Local Similarity: 36.1% Mismatches: 40
Query Match: 15.3% Indels: 22
Db: 5 Gaps: 4

US-10-071-510A-16 (1-493) x US-10-739-930-10816 (1-238)
QY 273 CCATTTTCTCCCGAATCCATTCCTCGAGGACGATTTGTGATGAGGCTCA 214
Db 56 ProAlaGlyAsnProProProAlaProSerAlaProProAlaProThrThProProGln 75
QY 213 GCAGACTTCTTCCAGCAGCC-----TCAAGGCCCGCGCAGCAGCTGTGCA 169
Db 76 AlaProGlnAlaProAlaThProProProAlaProGlnAlaProAlaThProProPro 95
QY 168 GCTCCCTGAGCTCTCTGCA----- 148
Db 96 AlaPro***ThProProAlaThProProProAlaProThThProProProAlaPro 115
QY 147 CCAGCGGACGACCCCTCCGAGAAAGACTTCCATCAACGACGCGCTCGCTTCACCA 88
Db 116 ProThGlnProProProAlaProThThThProProSerProProAlaThProProPro 135
QY 87 GGGACAGCTGG-----CCTCTTCTCCGGAATTCGACACGACGCTCAAACTCG 34
Db 136 ProAlaProAlaThThProProProSerProPrometSerProProProAlaThProPro 155
```

```
QY 33 CCTTTGGGAGCTGGGATCCCCG 10
Db 156 PrometAlaThr-----ProPro 161

RESULT 11
US-10-425-115-263989
; Sequence 263989, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425, 115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 263989
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(239)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_172373C.1.pdp
US-10-425-115-263989

Alignment Scores:
Pred. No.: 0.00599 Length: 239
Score: 141.00 Matches: 39
Percent Similarity: 35.3% Conservative: 10
Best Local Similarity: 28.1% Mismatches: 46
Query Match: 15.3% Indels: 44
Db: 4 Gaps: 4

US-10-071-510A-16 (1-493) x US-10-425-115-263989 (1-239)
QY 411 CCCCTAACCCACTTGTGACCTGGGCGCTGCGCCCTTCTCCACAGCCCTGTAGACAGC 352
Db 69 ProAlaProThAlaThProProProAlaProThThThProProProAlaProThr--- 87
QY 351 TCACGCGTCGACGATCCTGGGAATGAGATCCATGGAATGAGAAATCTGACTTGG 292
Db 87 ----- 87
QY 291 GGATGTTGTGTGAACCAATTTCTCCCGAATCCATTCCTCTCGACATGCC 232
Db 88 -----ThProProProAlaProThThThProProPro 98
QY 231 AGTTCTGATGAGGCTCAGCAGACTTCTTCACGACGCTCAAGGCCGACGACTGTG 172
Db 99 Ala-----ProThThThProProProAlaProThThTh 109
QY 171 CCAGCTCCCTGAGCTCTCTGACAGCAGGACGACCCCTCCGAGAAAGACTTTCATCA 112
Db 110 ProProProSerProProAlaSerPro-----ProProAlaProThThThProPro 126
QY 111 CCAGCAGCCCTGCGCTTCACACGAGGACAGCTGG-----CCTCTTCTCCGGAAT 58
Db 127 ProSerProProAlaSerProProProAlaProAlaThThProProProSerProPromet 146
QY 57 CTGCCACGAGCCTCTCAAACTCGGCTCTTGGACTGGGATCCCGCGGCGCGCG 1
Db 147 AlaProProAlaThThProProProProAlaThThProProProProAlaAlaPro 165

RESULT 12
US-10-767-701-40674
; Sequence 40674, Application US/10767701
```



```

; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40674
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C80510_1.pep
; US-10-767-701-40674

Alignment Scores:
Pred. No.: 0.00982 Length: 155
Score: 138.00 Matches: 36
Percent Similarity: 38.5% Conservative: 48
Best Local Similarity: 34.6% Mismatches: 4
Query Match: 15.0% Indels: 16
DB: 4 Gaps: 2

US-10-071-510A-16 (1-493) x US-10-767-701-40674 (1-155)
QY 282 TGGTGAACCATTTTCTTCCCGAATCCACTTCCTTCAGATGCCAATTTCTGA 223
DB 7 TTPProArgProLeuSerSerProArgAlaProSerProProAla----- 23
QY 222 TGAAGCTCAGCACTTTCTTCAGACAGCTCAAGCCCGCCAGCACTCGCAGCTCC 163
DB 24 -----AlaProProAlaProProProProProProAlaProProProPro 40
QY 162 TGAGCTCTCTCTGCACACCG-----CAGCACTT 133
DB 41 LeuAlaProProAlaProArgGlyAlaArgArgProArgGlyArgProProAlaPro 60
QY 132 CCGGAGAAAGCTTCTCCATCAGCAGCCCTGCGCTTCCACCGAGGAGAGCTGGCCT 73
DB 61 ProAlaGlyArgAlaAlaAlaProAlaProAlaArgProProAlaGlyArg 80
QY 72 CCTTCCCGGAGATTCGCGACAGCCTCAACTCGGCTTGGAGTCGGCATCCC 13
DB 81 ProGlyProGlyAlaAlaProGlyProProProSerArgAlaAlaProThrArgArgPro 100
QY 12 CCGGCGCGGCG 1
DB 101 ArgAlaProPro 104

RESULT 13
US-10-767-701-36014
; Sequence 36014, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36014
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure

```

```

; LOCATION: (1)..(189)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C79158_1.pep
; US-10-767-701-36014

Alignment Scores:
Pred. No.: 0.0101 Length: 189
Score: 138.00 Matches: 37
Percent Similarity: 42.3% Conservative: 4
Best Local Similarity: 38.1% Mismatches: 34
Query Match: 15.0% Indels: 22
DB: 4 Gaps: 4

US-10-071-510A-16 (1-493) x US-10-767-701-36014 (1-189)
QY 267 TCTCCCGAATTCATTCATCTCTGAGATGCGAGTTTGTATGAGCTCAGCAGAC 208
DB 3 AlaSerProSerProLeu----- 8
QY 207 TTTCTTCAGCAGCTCA-----AGCCCGCCAGCACTTGCAGCTCCCTGA 160
DB 9 LeuProProProAlaSerArgArgProProArgProValProProProProProPro 28
QY 159 GCTCTCTCGACCAAGGAGCAGCAGCCTCGGAGAGACTTTCATCAGCA-----GCC 106
DB 29 ProProProSerProArgSerProAlaProSerPro**ProProSerProProArgPro 48
QY 105 AGCCTGCGCTTCCAGCAGGAGCAGCTGGGCTCTTCTCCG-----GGAATTCGCA 52
DB 49 ProProAlaSerProProProAlaSerGlyProProSerProValLeuGlyArgArgPro 68
QY 51 CCAGCCTCTCAACTCGGCTTCTGGAGCTCGGCATCCCCGGGCGGCG 1
DB 69 ProValArgAlaCysArgProCysArgGlyProAlaProAlaProArgPro 85

RESULT 14
US-10-369-493-3962
; Sequence 3962, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3962
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Neurospora crassa
; US-10-369-493-3962

Alignment Scores:
Pred. No.: 0.0124 Length: 639
Score: 138.00 Matches: 39
Percent Similarity: 43.8% Conservative: 3
Best Local Similarity: 40.6% Mismatches: 36
Query Match: 15.0% Indels: 18
DB: 4 Gaps: 4

US-10-071-510A-16 (1-493) x US-10-369-493-3962 (1-639)
QY 255 CCACTTCATCTCTGAGATGCGAGTTTGTATGAGCTCAGCAGACTTCTT---CCA 199
DB 423 ProLeuProProGlyAlaPro-----GlyProAlaProProProLeuProPro 437

```

```
QY 198 GCAGCCTCAAGCGCCCGCAGACTTCGCCAGCTCCCTGAGCTCTCTGCGACCGGCG 139
DB 438 AAlaSerSerArgProProProMetLeuProThr-----ArgSerProAlaProPro 455
QY 138 CAGCCCTCGGAGAGAACTTCTCCATCAGCAGCGCCCTGCGGCTTCA----- 91
DB 456 AAlaProProLeuProThrSerAsnAlaProProProProProAlaThrGlnAla 475
QY 90 -----CCAGGAGCAGCTGGGCTCTCTTCGCGGAAATTCTGCCACCA 49
DB 476 ProProProProProLeuProAlaThrSerAlaProProProProProProAlaProPro 495
QY 48 GCCTCTCAAACTGGGCTCTTGGGACTCGGCATCCCGGCGCGCG 1
DB 496 AAlaProProAlaProProLeuProAlaAlaAlaProProProPro 511

RESULT 15
US-10-425-115-229737
; Sequence 229737, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 229737
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(180)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_141114C.1.pep
US-10-425-115-229737

Alignment Scores:
Pred. No.: 0.011 Length: 180
Score: 137.50 Matches: 38
Percent Similarity: 31.3% Conservative: 3
Best Local Similarity: 29.0% Mismatches: 39
Query Match: 14.9% Indels: 51
DB: 4 Gaps: 5

US-10-071-510A-16 (1-493) x US-10-425-115-229737 (1-180)
QY 393 CTTGGGGCTGGCCCTTCTCCAGCAGCCGTGACAGACTCAGCGTCAGCATGCC 334
DB 1 ProAlaHisIstPrProLeuLeuProProCys----- 11
QY 333 TGGGAATAGATCCATGGATGTGATGAATACTGACTTTGGGATGTTGGTGAAAA 274
DB 11 ----- 11
QY 273 CCATTTTCTCCCGAATCCACTTCATCTCTGCAGATGCCAGTTTGTGATGAGGCTCA 214
DB 12 -----**HiserProLeuProPro----- 18
QY 213 GGAAGATTCTTCCAGCAGCTCAAGCCCGCAGACTCTGCGAGCTCCCTGAGCTCCT 154
DB 19 -----Pro**ProArgArgProProGlnPro**Pro**Pro-----Pro 32
QY 153 CTTGACACAGCGCAGCAGCTCCGAGAGACTTCTCATACAGCCAGCCCTGGGCTT 94
DB 33 ProCysProProGlnProPro-----ArgAspSerProThrProProProThrPro 51
```

```
QY 93 CCACCAAGGAGCAAGCTGGGCTCTTCTCGGGAATTTGCCACAGCCTCTCAAACTCGG 34
DB 52 ProProProProHisProProProProProProHisProProProSerProProAsn 71
QY 33 CCTCTTGGGACTCGGCATCCCGGCGCGCGCG 1
DB 72 Pro-----ProProHisProProProProPro 80
```

Search completed: April 29, 2006, 04:09:01  
Job time : 124 secs



	Db	23	ProThrSerProProThralaThrProAlaProProThProThr-----	37
	Qy	345	GTGCAGCATCCTGGGAATGATCCATCCATGGATTGATGAGAAATCTCTACTTTGGGATGT	28
	Db	38	-----ThrPro-----	39
	Qy	285	TGTTGTGAAAAACCATTTCTTCCCGAATCCACTTCATCCTGTGAGATGCCAGTTTC	22
	Db	40	-----ProProAlaAlaIatnrProProProAl-----	48
	Qy	225	TGATGAGGCTCAGACGACTTTCTTCAGCAGCCTCAAGGCCGCGCACGACTGTGCCAGCT	16
	Db	49	-----SerAlaProProProValnrThrThrsSerProPro	59
	Qy	165	CCC-----TGAGCTCTCTGTGCACACGCGCAGACCCCTCCGAGAACTTTCATCA	11
	Db	60	ProValnrThrThralaProPro---ProAlaAsnProProProProValnrSerSerProPro	78
	Qy	111	CCAGCCAGC-----CTGCGGCTTCCACCCAGGGACAGCTGGGCGTCTCTTCGCGGAATT	58
	Db	79	ProIaIaSerProProProAlaIatnrProProProValAlaIaSerProProProProProValAla	98
	Qy	57	CTGCACACAGCCTCTCAACTCGGCGCTTTGGGACTCGGGATCCCCGGGGCGGCG	1
	Db	99	SerProProProAlaIatnrProProProValAlaIatnr-----ProProProAlaPro	115

```

RESULT 2
US-11-036-256-89
; Sequence 89, Application US/11036256
; Publication No. US20060026719A1
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA
; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; TITLE OF INVENTION: PEPTIDES/PROTEINS PRODUCED THEREBY
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; PRIOR FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 89
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid construct
US-11-036-256-89

Alignment Scores:
Pred. No.: 0.00271 Length: 411
Score: 128.50 Matches: 45
Percent Similarity: 35.0% Conservative: 5
Best Local Similarity: 31.5% Mismatches: 52
Query Match: 13.9% Indels: 41
DB: 7 Gaps: 7

US-10-071-510A-16 (1-493) x US-11-036-256-89 (1-411)
QY 411 CCCCTACCCACTCTTGACCTGAGGCGCTG-----CCCTTCTCCACAGCCC 364
||| ||||| ||||| |||||
Db 94 ProThrProThrProProProGlyProHisSerProProProProLeuSerProSerPro 113
||| ||||| ||||| |||||
QY 363 TGTCTGACAGACTCAACGCTCGACAGATGCTGGGATAGGATCGCATGGATTGATGAGA 304
||| ||||| ||||| |||||
Db 114 ThrProThrPro----- 117

```

```

QY      303  ATCTGACTTTGGGATGTTGTTGGTGAAGAAACCATTTTCTTCCCGGAATCAGTCTTCATCC 244
           |||||
           |||||
Db      118  -----ProLeuGlyProHisSerProProProThrLeuSerProSerProThrProThr 135
           |||||
           |||||
           |||||
QY      243  TGTGCAGATGCACGTTCTGATGAGAGCTCAGACAGATTCTTCCAGCAGCCTCAAGGCC 184
           |||||
           |||||
           |||||
Db      136  ProPro-----ProGlyProHisSerPro 143
           |||||
           |||||
           |||||
QY      183  GCCACGACTCTG-----CCAGCTCCGTAGCTCCTCTGCACACGAGGACGACCT 133
           |||||
           |||||
           |||||
Db      144  ProProProLeuSerProSerProThrPro---ThrProProLeuGlyProHisSerPro 162
           |||||
           |||||
           |||||
QY      132  CCGGAGAAAGACTTCTTCATCACCAGCCAGCCTCGGCTTCCACACGAGGACAGTGGGGCT 73
           |||||
           |||||
           |||||
Db      163  ProProThrLeuSerProSerPro---ThrProThrProProGlyValProGlyPro 181
           |||||
           |||||
           |||||
QY      72  -----CCTTCTCCGGGAATTCGTGCACACAGCCTCAAACTCGGGCTTGTGGGACTCG 19
           |||||
           |||||
           |||||
Db      182  HisSerProProProProLeuSerProSerProThrProThrProProLeuGlyProHis 201
           |||||
           |||||
           |||||
QY      18  CATCCCCCG 10
           |||||
           |||||
           |||||
Db      202  SerProPro 204

```

```

RESULT 3
US-11-232-440-19
; Sequence 19, Application US/11232440
; Publication No. US2006006843A1
; GENERAL INFORMATION:
; APPLICANT: STOECKER, JAY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING CANCER USING
; TITLE OF INVENTION: COMPONENTS OF THE U2 SPLICEOSOMAL PARTICLE
; FILE REFERENCE: MFP-031
; CURRENT APPLICATION NUMBER: US/11/232,440
; CURRENT FILING DATE: 2005-09-21
; PRIOR APPLICATION NUMBER: 60/612,310
; PRIOR FILING DATE: 2004-09-21
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: Patentin Ver. 3.3
; SEQ ID NO 19
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-232-440-19

Alignment Scores:
Pred. No.:      0.00404    Length:     464
Score:          126..50    Matches:   52
Percent Similarity:       38.8% Conservative: 12
Best Local Similarity:     31.5% Mismatches:   66
Query Match:           13.7% Indels:         35
DB:                      Gaps:              9

US-10-071-510A-16 (1-493) x US-11-232-440-19 (1-464)

Oy        440  CCAGACGAGGGGCCCTGTGTCGAGAACCCCTCAACTTGAACCTTGCGGCCTTGC 381
||| ||| ||| ::||| ||| ||||| 
Db        265  ProProthrnglyProlalProserdlYproProglYproProglmleuproProProla 284
||||| 

Oy        380  CCTTTTCCCACAGCCCTGTCTAGACAAGCTCACGGGTGACGATGCTGGAAATGAGATC 321
||| ||| ||||| ||| |:::||| 
Db        285  ProglyValHisProProProlalProvalValHisPro-----ProlasegrylVal 301
||| ||| ||||| ||| |:::||| 

Oy        320  CAT-----GGATTGATAGAANAATCTGACTTTGGATGTTSTGTGAATAAC 273
||| ||| :||| :||| :||| ||| :||| 
Db        302  HisProProlalProglyValHisProProla-----ProglyValHis 316
||||| 

Oy        272  CATTTCCTCCCGCATCCAC-----TTCCATCTCTGCGAGATGCCAGTT 228
||| ||| :||| ||||| ||| ||| 
Db        317  ProProlalaproglyValHisProProThrsergylValHisProProlalProglyVal 336
||||| 

Oy        227  TCTGATGAGGCTCAGCAGACTTTCTTCACAGCAGCTCAAAGCCCCGCCACGACTTCGCCAG 168
```

```
DB 337 HisProProAlaProGlyValHis--ProProAla-----ProGlyValHisProProAla 354
QY 167 CTCCTGAGCT---CCTCTGACCCAGCCAGCAGCCCT-----C 132
DB 354 IaProGlyValHisProProAlaProGlyValHisProProSerAlaGlyValHisP 374
QY 131 CGGAGAGACTTCTCCATCAGCCAGCCAGCCCTGCGCTTCCAGGAGAGAGCTGGG---- 76
DB 374 roGlnAlaProGlyValHisProAlaAlaProAlaValHisProGlnAlaProGlyValH 394
QY 75 --CTCTCTTCTCCGGGAATCTTCCACCAAGCCCTTCAAACTCGGCTCTTGGAGCTGGG 18
DB 394 IaProProAlaProGlyValHisProGlnAlaProGlyValHisPro-----G 410
QY 17 ATCCCCGGGCG 7
DB 410 InProProGly 413

RESULT 4
US-11-232-440-67
; Sequence 67, Application US/11232440
; Publication No. US2006006843A1
; GENERAL INFORMATION:
; APPLICANT: STOECKER, JAY
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING CANCER USING
; TITLE OF INVENTION: COMPONENTS OF THE U2 SPLICEOSOMAL PARTICLE
; FILE REFERENCE: MTP-031
; CURRENT APPLICATION NUMBER: US/11/232,440
; CURRENT FILING DATE: 2005-09-21
; PRIOR APPLICATION NUMBER: 60/612,310
; PRIOR FILING DATE: 2004-09-21
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 67
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-232-440-67

Alignment Scores:
Pred. No.: 0.00404 Length: 464
Score: 126.50 Matches: 52
Percent Similarity: 38.8% Conservative: 12
Best Local Similarity: 31.5% Mismatches: 66
Query Match: 13.7% Indels: 35
DB: 7 Gaps: 9

US-10-071-510A-16 (1-493) x US-11-232-440-67 (1-464)
QY 440 CCCAGAGAGGAGGGGCTGTGAGAGACCCCTTACCCACTTGTGACCTGGGCGCTGGC 381
DB 265 ProProThriProAlaProSerGlyProProGlyProProGlnLeuProProProAla 284
QY 380 CCCTTCTCCACAGCCCTGTAGACAGACTCAGCCGCTGAGATGCTGGGATAGATC 321
DB 285 ProGlyValHisProProAlaProAlaValHisPro-----ProAlaSerGlyVal 301
QY 320 CAT-----GGGATGATGAGAAATCTGACCTTGGGATGTTGTGGTGAAC 273
DB 302 HisProProAlaProGlyValHisProProAla-----ProGlyValHis 316
QY 272 CATTTTCTCCGGAATCCAC-----TTCCATCTCTGACAGATCCAGTT 228
DB 317 ProProAlaProGlyValHisProProThriSerGlyValHisProProAlaProGlyVal 336
QY 227 TCTGATGAGCTCAGCACTTCTTCCAGCAGCCCTCAAGGCCGCCAGCAAGCTGCCAG 168
DB 337 HisProProAlaProGlyValHis--ProProAla-----ProGlyValHisProProAla 354
QY 167 CTCCTGAGCT---CTCTGACAGCAGGAGCAGCCCT-----C 132
DB 354 IaProGlyValHisProProAlaProGlyValHisProProProSerAlaGlyValHisP 374
```

```
QY 131 CGGAGAGACTTCTCCATCAGCCAGCCCTTCCAGGAGAGAGCTGGG---- 76
DB 374 roGlnAlaProGlyValHisProAlaAlaProAlaValHisProGlnAlaProGlyValH 394
QY 75 --CTCTCTTCTCCGGGAATCTTCCACCAAGCCCTTCAAACTCGGCTCTTGGAGCTGGG 18
DB 394 IaProProAlaProGlyValHisProGlnAlaProGlyValHisPro-----G 410
QY 17 ATCCCCGGGCG 7
DB 410 InProProGly 413

RESULT 5
US-11-087-099-7887
; Sequence 7887, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7887
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Brassica napus
US-11-087-099-7887

Alignment Scores:
Pred. No.: 0.00613 Length: 647
Score: 124.50 Matches: 36
Percent Similarity: 48.4% Conservative: 10
Best Local Similarity: 37.9% Mismatches: 40
Query Match: 13.5% Indels: 9
DB: 7 Gaps: 6

US-10-071-510A-16 (1-493) x US-11-087-099-7887 (1-647)
QY 273 CCATTTCTTCCCGAATCCA-----CTTCACTCTTGCAGATGCCAGTTTCTGATGA 220
DB 35 ProSerSerProProProProProSerThriLeProThriSerProProSerSer----- 52
QY 219 GGCTCAGCAGACTTTCTTCCAGCAGCTCAAGCCCGCCAGCACTGCGAGCTCCCTGA 160
DB 53 ArgSerThriProSerAlaProProProSerProProThriProSerThriProGlySer-- 71
QY 159 GCTCTCTCTCCAGCAGGAGCAGCCCTCCGGAGAGACTTCTCA---TCACAGCCAGC 103
DB 72 ProProProLeuProGlnProProSerProProAlaProThriProGlySerProProAla 91
QY 102 CTGAGCTTCCACAGGAGCAGCTGGGCTCTTCT---CCGGGAATTTGCCAGCAGCC 46
DB 92 ProValThriProProThriArgAnProProProSerAlaProGly-----ProProSer 109
QY 45 TCTCAACTCGGCTCTTGGAGCTCGGCAATCCCGGCGCGCGC 1
DB 110 AsnProSerArgGlnGlyGlySerProArgProProSerSerPro 124

RESULT 6
US-11-188-298-18316
; Sequence 18316, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
```

```

; SEQ ID NO 18316
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Brassica napus
US-11-188-298-18316

Alignment Scores:
Pred. No.: 0.00613 Length: 647
Score: 124.50 Matches: 36
Percent Similarity: 48.4% Conservative: 10
Best Local Similarity: 37.9% Mismatches: 40
Query Match: 13.5% Indels: 9
DB: 7 Gaps: 6

US-10-071-510A-16 (1-493) x US-11-188-298-18316 (1-647)
QY 273 CCATTTTCTTCCCGAATCCA-----CTTCCATCTCTGACAGATCCAGTTTCTGATGA 220
DB 35 ProSerSerProProProProSerThrIleProThnSerProProProSer----- 52
QY 219 GGCTCAGACAGACTTCTTCCAGACGCTTAAGGCCCGCCACACACTCTGCCAGCTCCCTGA 160
DB 53 ArgSerThrProSerAlaProProProSerProProThrProSerThrProGlySer--- 71
QY 159 GGTCTCTCCGACACGCGACGACACCTCCGGAAGAAGACTTCTCA---TCACAGACGAC 103
DB 72 ProProProLeuProGlnProSerProProAlaProThnThrProGlySerProProAla 91
QY 102 CCTGCGCTTCCACACGAGACAGCTGGGCTCTCT---CCGGAATTCTGCCACAGCC 46
DB 92 ProValThrProProThrArgAnProProProSerValProGly-----ProProSer 109
QY 45 TCTCAAACTCGCGCTCTTGGGACTCGGATCCCGCGCGCGCGC 1
DB 110 AsnProSerArgGluGlyGlySerProArgProProSerSerPro 124

RESULT 7
US-11-036-256-21
; Sequence 21, Application US/11036256
; Publication No. US20060026719A1
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA
; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; PRIOR FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 21
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid construct
US-11-036-256-21

Alignment Scores:
Pred. No.: 0.00742 Length: 183
Score: 123.00 Matches: 44
Percent Similarity: 34.8% Conservative: 5
Best Local Similarity: 31.2% Mismatches: 52
Query Match: 13.3% Indels: 40
DB: 7 Gaps: 7

```

```

US-10-071-510A-16 (1-493) x US-11-036-256-21 (1-183)
QY 411 CCCCTAACCCACTCTTGAACCTGGGGCTTG-----CCGCTTCTCCACAGCCC 364
DB 56 ProThrProThrProProProGlyProHisSerProProProLeuSerProSerPro 75
QY 363 TGCTAGACAGACTCAGCGGTGACAGATCGTGGAAATGATCCATGGAGATTGATGAGA 304
DB 76 ThrProThrPro----- 79
QY 303 ATCTGACTTGGAGATGTTGTTGGTAACAACATTTCTTCCCGAATTCACCTTCATCC 244
DB 80 -----ProLeuGlyProHisSerProProProThrLeuSerProSerProThrProThr 97
QY 243 TCTGAGATGCCAGTTTCTGATGAGGCTCAGAGACTTCTTCCAGCAGCCCAAGGCC 184
DB 98 ProPro-----ProGlyProHisSerPro 105
QY 183 GCCACGACTTG-----CAGCTCTGAGCTCTCTGACACGACGACGACCT 133
DB 106 ProProProLeuSerProSerProThrPro---ThrProProLeuGlyProHisSerPro 124
QY 132 CCGGAAAGACTTCTCATCACACGACGCTGCGCTTCCACAGGAGACAGTGGGCTT 73
DB 125 ProProThrLeuSerProSerPro---ThrProThrProProProGlyProHisSerPro 143
QY 72 CCTTCTCCGGAATTCTCCACACGACCTCTCAAACTCGGCTTCTGGGACTCGGATCCC 13
DB 144 ProProPro---LeuSerProSerProThrProThrProProLeuGlyProHisSerPro 162
QY 12 CCG 10
DB 163 Pro 163

RESULT 8
US-11-036-256-31
; Sequence 31, Application US/11036256
; Publication No. US20060026719A1
; GENERAL INFORMATION:
; APPLICANT: KIELISZEWSKI, MARCIA
; APPLICANT: XU, JIANFENG
; TITLE OF INVENTION: METHODS OF PRODUCING PEPTIDES/PROTEINS IN PLANTS AND
; FILE REFERENCE: 27211/04130
; CURRENT APPLICATION NUMBER: US/11/036,256
; PRIOR FILING DATE: 2005-01-14
; PRIOR APPLICATION NUMBER: 60/602,562
; PRIOR FILING DATE: 2004-08-18
; PRIOR APPLICATION NUMBER: 60/582,027
; PRIOR FILING DATE: 2004-06-22
; PRIOR APPLICATION NUMBER: 60/536,486
; PRIOR FILING DATE: 2004-01-14
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 31
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: amino acid construct
US-11-036-256-31

Alignment Scores:
Pred. No.: 0.00755 Length: 228
Score: 123.00 Matches: 44
Percent Similarity: 34.8% Conservative: 5
Best Local Similarity: 31.2% Mismatches: 52
Query Match: 13.3% Indels: 40
DB: 7 Gaps: 7
US-10-071-510A-16 (1-493) x US-11-036-256-31 (1-228)

```

Qy	411	CCSCCTTCCCAAGCTCTTGA	CTGGGGCTGG-----CCSCCTTCCCAAGCTCC	364
Db	56	ProThrProThrProProGlyProHisSerProProProLeuSerProSerPro		75
Qy	363	TGCTACACAGACTCAGCGCTGACGATGCTGGGAATGATTCATGGAGATTGATGAA		304
Db	76	ThrProThrPro-----		79
Qy	303	ATCCTGACTTTGGGATGTGTGGTGA	AAACCATTTTCTCCCGGATCCACTTCCATCC	244
Db	80	-----ProLeuGlyProHisSerProProProThrLeuSerProSerProThrProThr		97
Qy	243	TCTGCAGATGCCAGTGTTCGATGAGAGCTCAGACACTTCTTCCAGACGCTCAAGGCC		184
Db	98	ProPro-----	-----ProGlyProHisSerPro	105
Qy	183	GCCAGCACTCTG-----CCAGCTCCCTGAGTCTCTCTGCAACAGGACACACT		133
Db	106	ProProProLeuSerProSerProThrPro--ThrProProLeuGlyProHisSerPro		124
Qy	132	CCGGAGAAAGACTTCTGCATCACCAGACGACCTCGGCTTCACACAGGAGACGTGGGCT		73
Db	125	ProProThrLeuSerProSerPro--ThrProThrProProProGlyProHisSerPro		143
Qy	72	CCTTTCCGGGAATTCTGCACAGGCTCTCAACTCGAGCTCTTGGGACTCGGCAATCC		13
			:::	
Db	144	ProProPro--LeuSerProSerProThrProThrProProLeuGlyProHisSerPro		162
Qy	12	CCG 10		
Db	163	Pro 163		

```
QY 87 GGGAGAGTGGGCTCTTCT-----CGGGAAATTCTGCCACGACTTCAAACTCGG 34
DB 95 ProSeSeProProProAlaSerProProProSeSeSeProProProSeSeProPro 114
QY 33 CCTCTTGGAGCTCGCATCCCGGCGCGCG 1
DB 115 ProPhSeSePro-----ProProAlaThrPro 123

RESULT 11
US-11-188-298-694
; Sequence 694, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; PRIOR FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 694
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Triticum aestivum
; NAME/KEY: unsure
; LOCATION: (1)..(243)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-188-298-694

Alignment Scores:
Pred. No.: 0.00837 Length: 243
Score: 122.50 Matches: 41
Percent Similarity: 31.7% Conservative: 36
Best Local Similarity: 29.5% Mismatches: 26
Query Match: 13.3% Indels: 69
Gaps: 7

US-10-071-510A-16 (1-493) x US-11-188-298-694 (1-243)
QY 417 TGGAGAGCCCTTACCCAGACTCTTGTGAGGCGCTGAGCCCTTCTCCAGAGCCCTGCTAG 358
DB 172 TrrProProLeuProThrPro-----LeuPheProThrPro----- 183
QY 357 ACAGACTCAAGCGCTGCAGCATGCTGGAAATGATTCATGGATTGATGAGAAATCTTG 298
DB 183 ----- 183
QY 297 ACTTTGGAGATGTTGTGGTGAACATTTTCTTCCCGAATCCACTTTCATCTCTGCA 238
DB 184 -----ProLysProSeSeProAsnProPhe----- 191
QY 237 GATGCCAGTTTCTGATGAGGCTCAGCAGACTTTCTTCCAGAGCTCAAGGCCCGCCAGC 178
DB 192 -----PheLeuProLeu-----LeuThr 197
QY 177 ACTTGCAGCTCCCTGAGACTCTCTGACACAGCGAGACCCCTCGGAGAAAGACTTCT 118
DB 198 HsLeuProProPro--ProProProProProProTyHisProProPheProThrAsn 216
QY 117 CCATCAACGAGCCCTGCGCTTCCACAGGAGACAGCTGGGCTCTCTTCTCCGGGAATT 58
DB 217 ProProProSeSeLeuProProProPro-----ThrProProSeSeProHisAs 233
QY 57 CTGCCACGAGCCTCAAACTGGCGCTTTGGAGACTCGGATCCCGCGGCGCGCGCG 1
DB 234 ProProPro-----ProProProGlnPro 241

RESULT 12
US-11-150-845-12
; Sequence 12, Application US/11150845
```

```
; Publication No. US20060003399A1
; GENERAL INFORMATION:
; APPLICANT: Cytokine, Inc.
; APPLICANT: Tomasevic, Nenad.
; APPLICANT: Jia, Zhiheng
; APPLICANT: Sakowicz, Roman
; APPLICANT: Pierce, Daniel
; APPLICANT: Finer, Jeffrey
; TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY
; FILE REFERENCE: 020552-007720US
; CURRENT APPLICATION NUMBER: US/11/150,845
; PRIOR FILING DATE: 2005-06-10
; PRIOR APPLICATION NUMBER: US 60/673,444
; PRIOR FILING DATE: 2005-04-20
; PRIOR APPLICATION NUMBER: 60/578,949
; PRIOR FILING DATE: 2004-06-10
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-150-845-12

Alignment Scores:
Pred. No.: 0.0129 Length: 408
Score: 120.50 Matches: 47
Percent Similarity: 33.9% Conservative: 9
Best Local Similarity: 28.5% Mismatches: 50
Query Match: 13.1% Indels: 59
Gaps: 7

US-10-071-510A-16 (1-493) x US-11-150-845-12 (1-408)
QY 467 CCAAGTCTTCAGAGAGCATGTTGCCCCAGACAGGAGGGGCTGTGCTGGAGACCC 408
DB 182 ProProProProSeSeArg-----GlyGlyProProProProPro 194
QY 407 TACCCAGCTTTGAGCTGGGGCTGGCCCTTCTCCACAGCCCTGTAGACAGACTCAC 348
DB 195 ProProProHisAsnSerGlyProProPro-----ProProAlaHisGlyArgGly 211
QY 347 GCGTGAGAGTCCCTGGGAATGATTCATGGATTGATGAGAAATCTGACTTTGGAT 288
DB 212 AlaProProProProProSeSeArg--Ala----- 220
QY 287 GTTGTGGTAAACCATTTTCTTCCCGAATCCATTCCATCTCTGCAATGCCAGTT 228
DB 221 -----ProThrAlaAlaProProPro----- 227
QY 227 TCTGATGAGGCTCAGACAGACTTCTTCCAGAGCCTCAAGGCCCGCAGACTTGCCAG 168
DB 228 -----ProProProSeSeArgProSeSeValAlaValPro- 238
QY 167 CTCCTGAGCTCTCTCTGACACAGCAGACCTTCGAGAAAGACTTCTCCATCAGCAG 108
DB 229 -----ProProProProAsnArgMetGlyProPro-----ProProProA 252
QY 107 CCAGCCCTGCGCTTCCACAGGAGACAGCTGGGCTCTCTTCTCCGGGAATTCTG----- 55
DB 252 lAlaLeuProSeSeSeValProProProProProProProProSeSeValLeuGlyValG 272
QY 54 -----CCACAGGCTCTCAAACTGGCGCTTGGGAGCTGGGATCCCG 12
DB 272 lProValAlaProProProProProProProProProProProProProProProPro 292
QY 11 CGGGCGGCGCG 1
DB 292 roGlyLeuPro 295

RESULT 13
US-11-150-487-12
; Sequence 12, Application US/11150487
```



Publication No. US20060024786A1  
 GENERAL INFORMATION:  
 APPLICANT: CytoGenetics, Inc.  
 APPLICANT: Tomasevic, Nenad  
 APPLICANT: Russell, Alan  
 APPLICANT: Wang, Maming  
 APPLICANT: Sakowicz, Roman  
 APPLICANT: Beraud, Christophe  
 TITLE OF INVENTION: WASP AND N-WASP CONSTRUCTS AND METHODS OF EXPRESSING SUCH  
 TITLE OF INVENTION: CONSTRUCTS  
 FILE REFERENCE: 020552-00791005  
 CURRENT APPLICATION NUMBER: US/11/150,487  
 PRIOR FILING DATE: 2005-06-10  
 PRIOR APPLICATION NUMBER: US 60/578,913  
 PRIOR FILING DATE: 2004-06-10  
 NUMBER OF SEQ ID NOS: 70  
 SOFTWARE: PatentIn version 3.3  
 SEQ ID NO 12  
 LENGTH: 408  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc.feature  
 LOCATION: (1)..(408)  
 OTHER INFORMATION: 98N-WASP  
 US-11-150-487-12

Alignment Scores:  
 Pred. No.: 0.0129 Length: 408  
 Score: 120.50 Matches: 47  
 Percent Similarity: 33.9% Conservative: 9  
 Best Local Similarity: 28.5% Mismatches: 50  
 Query Match: 13.1% Indels: 59  
 DB: 7 Gaps: 7

US-10-071-510a-16 (1-493) x US-11-150-487-12 (1-408)

QY 467 CCAAGTCTCAGACGACGATGTTGCCAGACAGGAGGCGCTGTGCTGAGACCCC 408  
 DB 182 ProProProProProSerArg-----GlyGlyProProProProProPro 194  
 QY 407 TACCCACTTGTGACCTGGGCGCTGCGCTTCTCCACAGCCCTGTAGACAGACTCAC 348  
 DB 195 ProProProHisHisSerGlyProProPro-----ProProAlaArgGlyArgGly 211  
 QY 347 GCCTGCACGATGCTGGGAATAGATCCATGATGATGATGATGATGATGATGATGAT 288  
 DB 212 AlaProProProProProSerArg--Ala----- 220  
 QY 287 GTTGTGTGTAACCATTTTCTTCCCGCAATCCATTCCTCTGCAGATGCCAGTT 228  
 DB 221 -----ProThrAlaAlaProProPro----- 227  
 QY 227 TCTGATGAGCTCAGACACTTCTTCCAGCAGCCTCAAGGCCGCGCAGACTCTGCCAG 168  
 DB 228 -----ProProProSerArgProSerValAlaValPro- 238  
 QY 167 CTCCTGAGCTCTCTGACACGAGCAGCAGCCTCCGGAGAAAGACTTCTCCATCAGCAG 108  
 DB 239 -----ProProProProHisHisSerGlyProProPro-----ProProProA 252  
 QY 107 CCAAGCCTGGGCTTCCACGAGGAGAGCTGGGCTCTCTCCGGAATTCG----- 55  
 DB 252 lAlaProSerSerAlaProSerGlyProProProProProProProProProProPro 272  
 QY 54 -----CCACGAGCTCTCAAACTCGGCTCTTGGGACTCGGATCCCC 12  
 DB 272 lProValAlaProProProProProProProProProProProProProProProPro 292  
 QY 11 CGGGCCGGCCG 1  
 DB 292 roGlyLeuPro 295

RESULT 14

US-11-134-563-6  
 Sequence 6, Application US/11134563  
 Publication No. US20050287569A1

GENERAL INFORMATION:  
 APPLICANT: Leon, John M.  
 APPLICANT: Campellone, Kenneth G.  
 TITLE OF INVENTION: ESPRU NUCLEIC ACIDS AND PROTEINS AND  
 TITLE OF INVENTION: USES THEREOF  
 FILE REFERENCE: 07917-280001  
 CURRENT APPLICATION NUMBER: US/11/134,563  
 PRIOR FILING DATE: 2005-05-20  
 PRIOR APPLICATION NUMBER: US 60/573,600  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 6  
 LENGTH: 505  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-11-134-563-6

Alignment Scores:  
 Pred. No.: 0.0131 Length: 505  
 Score: 120.50 Matches: 47  
 Percent Similarity: 33.9% Conservative: 9  
 Best Local Similarity: 28.5% Mismatches: 50  
 Query Match: 13.1% Indels: 59  
 DB: 7 Gaps: 7

US-10-071-510a-16 (1-493) x US-11-134-563-6 (1-505)

QY 467 CCAAGTCTCAGACGACGATGTTGCCAGACAGGAGGCGCTGTGCTGAGACCCC 408  
 DB 279 ProProProProProSerArg-----GlyGlyProProProProProPro 291  
 QY 407 TACCCACTTGTGACCTGGGCGCTGCGCTTCTCCACAGCCCTGTAGACAGACTCAC 348  
 DB 292 ProProProHisHisSerGlyProProPro-----ProProAlaArgGlyArgGly 308  
 QY 347 GCCTGCACGATGCTGGGAATAGATCCATGATGATGATGATGATGATGATGATGAT 288  
 DB 309 AlaProProProProProSerArg--Ala----- 317  
 QY 287 GTTGTGTGTAACCATTTTCTTCCCGCAATCCATTCCTCTGCAGATGCCAGTT 228  
 DB 318 -----ProThrAlaAlaProProPro----- 324  
 QY 227 TCTGATGAGCTCAGACACTTCTTCCAGCAGCCTCAAGGCCGCGCAGACTCTGCCAG 168  
 DB 325 -----ProProProSerArgProSerValAlaValPro- 335  
 QY 167 CTCCTGAGCTCTCTGACACGAGCAGCAGCCTCCGGAGAAAGACTTCTCCATCAGCAG 108  
 DB 336 -----ProProProProHisHisSerGlyProProPro-----ProProProA 349  
 QY 107 CCAAGCCTGGGCTTCCACGAGGAGAGCTGGGCTCTCTCCGGAATTCG----- 55  
 DB 349 lAlaProSerSerAlaProSerGlyProProProProProProProProProProPro 369  
 QY 54 -----CCACGAGCTCTCAAACTCGGCTCTTGGGACTCGGATCCCC 12  
 DB 369 lProValAlaProProProProProProProProProProProProProProProPro 389  
 QY 11 CGGGCCGGCCG 1  
 DB 389 roGlyLeuPro 392

RESULT 15  
 US-11-150-845-4  
 Sequence 4, Application US/11150845  
 Publication No. US20060003399A1  
 GENERAL INFORMATION:

Search completed: April 29, 2006, 04:09:46  
Job time : 23 secs

APPLICANT: Cycokinetics, Inc.  
APPLICANT: Tomasevic, Nenad  
APPLICANT: Jia, Zhiheng  
APPLICANT: Sakowicz, Roman  
APPLICANT: Pierce, Daniel  
APPLICANT: Finer, Jeffrey  
TITLE OF INVENTION: HIGH THROUGHPUT ACTIN POLYMERIZATION ASSAY  
FILE REFERENCE: 020552-007720US  
CURRENT APPLICATION NUMBER: US/11/150,845  
CURRENT FILING DATE: 2005-06-10  
PRIOR APPLICATION NUMBER: US 60/673,444  
PRIOR FILING DATE: 2005-04-20  
PRIOR APPLICATION NUMBER: 60/578,949  
PRIOR FILING DATE: 2004-06-10  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 4  
LENGTH: 505  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(505)  
OTHER INFORMATION: FL-N-WASP  
US-11-150-845-4

Alignment Scores:  
Pred. No.: 0.0131 Length: 505  
Score: 120.50 Matches: 47  
Percent Similarity: 33.9% Conservative: 9  
Best Local Similarity: 28.5% Mismatches: 50  
Query Match: 13.1% Indels: 59  
DB: 7 Gaps: 7

US-10-071-510a-16 (1-493) x US-11-150-845-4 (1-505)

QY 467 CCNAGTCTCAGACGAGCATGTTGCCCGACAGAGGAGGCGCTGTGCTGGAGACCCC 408  
DB 279 ProProProProSerArg-----GlyGlyProProProProPro 291  
QY 407 TACCCACCTTTGACCTGGGCGCTGGCCCTTCTCCACAGCCCTGTAGACAGACTCAG 348  
DB 292 ProProProHisbSerGlyProProPro-----ProProHisGlyArgGly 308  
QY 347 GCGTCACGATGCTGGGATAGATCCATGGGATTGATGAGAAATCTGACTTTGGGAT 288  
DB 309 AlaProProProProSerArg--Ala----- 317  
QY 287 GTTGTGTGMAAACCATTTTCTCCCGAATCCACTTCATCTCTGCGATGCCAGTT 228  
DB 318 -----ProthAlaAlaProProPro----- 324  
QY 227 TCTGATAGAGCTCAGCAGACTTTCTTCAGACGCTCAAGGCCCGCACGACTCTGCCAG 168  
DB 325 -----ProProProSerArgProSerValAlaValPro- 335  
QY 167 CTCCTGAGCTCTCTCTGCACACGCGGACGACCTCCGAGAAAGACTTCTCCATCACAG 108  
DB 336 -----ProProProProAsnArgMetLysProPro-----ProProProA 349  
QY 107 CCAGCCCTGGGCTTCACGACGAGGAGAGCTGCGCTCTTCTCCGGGAAATTTCTG----- 55  
DB 349 lAlenProSerSerAlaProSerGlyProProProProProProProProProProProPro 369  
QY 54 -----CCACGACCTTCMAACTCGGCTCTTGGGACTCGGCATCCCC 12  
DB 369 lYProValAlaProProProProProProProProProProProProProProProProPro 389  
QY 11 CGGGCGGGCGG 1  
DB 389 roGlyLeuPro 392